

CRFE

SEARCH REQUEST FORM

Access DB#

84868

Scientific and Technical Information Center

Requester's Full Name: David Fox Examiner #: 65401 Date: 1/21/03
Art Unit: 11638 Phone Number 30 8-0280 Serial Number: 09/701,023
Mail Box and Bldg/Room Location: CM1 Results Format Preferred (circle): PAPER DISK E-MAIL
9E12 9E15 9E15

If more than one search is submitted, please prioritize searches in order of need. *mej*

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched, include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc. if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: _____ Point of Contact: _____
Inventors (please provide full names): _____ Susan Hanley
Technical Info. Specialist
CM1 8605 Tel: 305-4053

Earliest Priority Filing Date: _____ *received 2/2/03*

For Sequence Searches Only Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.

Please do sequence search for
SEQ ID NO: 4 (protein) and
SEQ ID NO: 1, nucleotides 81-1024
gene encoding it).

best 3/13/03

Amendment due 2/18/03

1-1302
4-~~000~~314 *THANK YOU*

David

STAFF USE ONLY

Type of Search	Vendors and cost where applicable
Searcher: <u>Hanley</u>	NA Sequence (#) <u>1</u> STN _____
Searcher Phone #: _____	AA Sequence (#) <u>1</u> Dialog _____
Searcher Location: _____	Structure (#) _____ Questel/Orbit _____
Date Searcher Picked Up: <u>1/22</u>	Bibliographic _____ Dr. Link _____
Date Completed: <u>1/22</u>	Litigation _____ Lexis/Nexis _____
Searcher Prep & Review Time: <u>2</u>	Fulltext _____ Sequence Systems <u>01</u>
Critical Prep Time: <u>3</u>	Patent Family _____ WWW/Internet _____
Online Time: <u>7</u>	Other _____ Other (specify) _____

This Page Blank (uspto)

score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

M nucleic - nucleic search, using sw model

on on: January 25, 2003, 21:54:27 ; Search time 2843 seconds
9563.401 Million cell updates/sec

title: US-09-701-023-1_COPY_81_1024
seq score: 1 tggagactctctctctc.....atttgctctgagactctaa 944
sequence:

coring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

searched: 2054640 seqs, 14551402878 residues

total number of hits satisfying chosen parameters: 4109280

inimum db seq length: 0

aximum db seq length: 2000000000

out-processing: Minimum Match 0%
Maximum Match 100%
Missing first 45 summaries

database:

- 1: gb.pat.*
- 2: gb.pat.*
- 3: gb.pat.*
- 4: gb.pat.*
- 5: gb.pat.*
- 6: gb.pat.*
- 7: gb.pat.*
- 8: gb.pat.*
- 9: gb.pat.*
- 10: gb.pat.*
- 11: gb.pat.*
- 12: gb.pat.*
- 13: gb.pat.*
- 14: gb.pat.*
- 15: gb.pat.*
- 16: gb.pat.*
- 17: gb.pat.*
- 18: gb.pat.*
- 19: gb.pat.*
- 20: gb.pat.*
- 21: gb.pat.*
- 22: gb.pat.*
- 23: gb.pat.*
- 24: gb.pat.*
- 25: gb.pat.*
- 26: gb.pat.*
- 27: gb.pat.*
- 28: gb.pat.*
- 29: gb.pat.*
- 30: gb.pat.*
- 31: gb.pat.*
- 32: gb.pat.*
- 33: gb.pat.*
- 34: gb.pat.*
- 35: gb.pat.*
- 36: gb.pat.*
- 37: gb.pat.*
- 38: gb.pat.*
- 39: gb.pat.*
- 40: gb.pat.*
- 41: gb.pat.*

SUMMARIES

Result	Score	Query	Match	Length	DB	ID	Description
1	944	100.0	1302	8	AF159255	AF159255	Arabidopsis thaliana
2	736	78.0	1133	8	AF146794	AF146794	Arabidopsis thaliana
3	732.8	77.6	82381	8	AF176719	AF176719	Arabidopsis thaliana
4	732.8	77.6	83371	8	AF146794	AF146794	Arabidopsis thaliana
5	45.8	4.9	13716	6	AF146794	AF146794	Arabidopsis thaliana
6	45.8	4.9	13716	6	AF146794	AF146794	Arabidopsis thaliana
7	41.8	4.4	141632	2	AC104900	AC104900	Mus musculus
8	41.8	4.4	139072	2	AF136531	AF136531	Human DNA
9	41.8	4.4	139072	2	AF136531	AF136531	Human DNA
10	41.8	4.4	139072	2	AF136531	AF136531	Human DNA
11	40.4	4.3	1472	1	AF159255	AF159255	Arabidopsis thaliana
12	40.4	4.3	1472	1	AF159255	AF159255	Arabidopsis thaliana
13	40.4	4.3	1472	1	AF159255	AF159255	Arabidopsis thaliana
14	40.4	4.3	1472	1	AF159255	AF159255	Arabidopsis thaliana
15	40.4	4.3	1472	1	AF159255	AF159255	Arabidopsis thaliana
16	40.4	4.3	1472	1	AF159255	AF159255	Arabidopsis thaliana
17	39.8	4.2	131903	9	HS360618	HS360618	Human DNA
18	39.8	4.2	139843	2	AF158445	AF158445	Human DNA
19	39.8	4.2	139843	2	AF158445	AF158445	Human DNA
20	39.8	4.2	139843	2	AF158445	AF158445	Human DNA
21	39.8	4.2	139843	2	AF158445	AF158445	Human DNA
22	39.8	4.2	139843	2	AF158445	AF158445	Human DNA
23	39.8	4.2	139843	2	AF158445	AF158445	Human DNA
24	39.8	4.2	139843	2	AF158445	AF158445	Human DNA
25	39.8	4.2	139843	2	AF158445	AF158445	Human DNA
26	39.8	4.2	139843	2	AF158445	AF158445	Human DNA
27	39.8	4.2	139843	2	AF158445	AF158445	Human DNA
28	39.8	4.2	139843	2	AF158445	AF158445	Human DNA
29	39.8	4.2	139843	2	AF158445	AF158445	Human DNA
30	39.8	4.2	139843	2	AF158445	AF158445	Human DNA
31	38.6	4.1	103675	9	AF136223	AF136223	Human DNA
32	38.6	4.1	147598	9	AC093503	AC093503	Human DNA
33	38.6	4.1	147598	9	AC093503	AC093503	Human DNA
34	38.6	4.1	147598	9	AC093503	AC093503	Human DNA
35	38.6	4.1	147598	9	AC093503	AC093503	Human DNA
36	38.6	4.1	147598	9	AC093503	AC093503	Human DNA
37	38.6	4.1	147598	9	AC093503	AC093503	Human DNA
38	38.6	4.1	147598	9	AC093503	AC093503	Human DNA
39	38.6	4.1	147598	9	AC093503	AC093503	Human DNA
40	38.6	4.1	147598	9	AC093503	AC093503	Human DNA
41	38.6	4.1	147598	9	AC093503	AC093503	Human DNA
42	38.6	4.1	147598	9	AC093503	AC093503	Human DNA
43	38.6	4.1	147598	9	AC093503	AC093503	Human DNA
44	38.6	4.1	147598	9	AC093503	AC093503	Human DNA
45	38.6	4.1	147598	9	AC093503	AC093503	Human DNA

ALIGNMENTS

RESULT 1	AF159255	1302 bp	linear	PLN 04-SEP-1999
DEFINITION	Arabidopsis thaliana spotyless (SP1) mRNA, complete cds.			
ACCESSION	AF159255			
VERSION	AF159255.1	GI:55662939		
ORGANISM	Arabidopsis thaliana			
REFERENCE	Bukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons: core eudicots; Rosidae: euRosidae(1); Brassicales; Brassicaceae; Arabidopsis. Yang W.C., Ye D., Xu J. and Sundaresan, V. The SPOTYLESS gene of Arabidopsis is required for initiation of			

Pred. No. is the number of results predicted by chance to have a


```
sporogenesis="
/code_start=1
/product="spores"
/db_xref="GI:5007009.1"
translation="MATELFNMFSTDCNSVNPENLRLNVLNWSGEARTFLASG
PMPNPNRNDKQKKRRLGQNNYNSGGSGPKRTTIPPMNGDOYLODNR
EYVDPVQVLFQSSSLASGSGVSVIWDPIVSENGVSTSTTHSISLN
PMPNPNRNDKQKKRRLGQNNYNSGGSGPKRTTIPPMNGDOYLODNR
EYVDPVQVLFQSSSLASGSGVSVIWDPIVSENGVSTSTTHSISLN
KEVETGPKGVNSVATSLASGSGCPTNLSL"

```

variation

gene- NZZ
/note="compared to Arabidopsis thaliana Columbia ecotype

variation

N

erv Match

Local SI

—

000007 11

[illegible]

62 TITIG

121 AGAGT

122 AGAGT

181 TGAGF

182 TCAG

717 70000

— 6 —

1111

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

11111

352 GCAAC

421 CTCCF

422 CTCCF

481 CTCAT

492 CTCAE

330

...

706

562 TAGT

[illegible]

[illegible]

/note="strong similarity to proline synthetase associated
- Homo sapiens, PID:d1037830

[illegible][illegible]

b 73768 AAAGGAAAGAAGGAAGCAAGGAAGGAAGGAAGGAAGAACCCACCTCAAAACGACTTA 73827
|||||
y 254 AGTCGGACACGCGTCAATGC 274
|||
b 73828 AGTCGAAGCAGGTATTGAAT 73848
|||||

RESULTS
AL136531/c
LOCUS
SYNPHALLIN
DEFINITION
b 73828 AGTCGAAGCAGGTATTGAAT 73848
|||||

CESSION
AL136531.16 GI:11139883
FTIG: CGC island; F506; FKBP1A; KIAA0078; KIAA0374; PPPI; RAD21;
scf fold protein; SCDFP; SYNTAC18; SYCL; syntaphillin;
human.
ORIGIN
Homo sapiens
ORGANISM
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catartia; Hominiidae; Homo.
REFERENCE
JOURNAL
TITLE
DIRECT SUBMISSION
Submitted (06-FEB-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: hunkquery@sanger.ac.uk Clone
ID: 11139883.1 TO 119072)
COMMENT
On Nov 13, 2000 this sequence version replaced gi:10045291.
During differences are found these are annotated as variations
with the original clone name. The overlapping clone name. Note that the
variation was observed at the end of the overlap. This variation
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
The following abbreviations are used to associate primary accession
numbers with their corresponding databases:
EMBL: EN, SW, SNRSPROT; TrEMBL: TR, SW, WMAPSP; Information
on the WormPe database can be found at
<http://www.sanger.ac.uk/Projects/C.elegans/WormPe> This sequence
was generated from part of bacterial clone contigs of human
chromosome 20. The clones were generated by Genomic
Mapping Group. Further information can be found at
<http://www.sanger.ac.uk/HGP/Chr20>
IMPORTANT: This sequence is not the entire insert of clone
AP11-314N3 it may be shorter because we sequence overlapping
sequences. The true right end of clone RP4-776P14 is at 118973 in this
sequence. This sequence was finished as follows unless otherwise
noted: all regions were either double-stranded or sequenced with an
average quality >= 30; an attempt was made to resolve all sequencing
problems, such as compressions and repeats; all regions were
covered by at least one plasmid subclone or more than one M13
subcloned and the assembly was confirmed by restriction digest.
AP11-314N3 was constructed by PCR constructed by the group
of Peter de Jong. For further details see
<http://www.chori.org/bacpac/home.htm>
VECTOR: pBAKc3.6.
LOCALIZATION: Qualifiers
Location: 119072
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="20"
/clone="AP11-314N3"
/contig="RP4-776P14-NCBI-11.2"
repeat_region
c . :184

/note=-L1 repeat: matches 2983..3159 of consensus"
 171..247
 repeat_region
 /note=-L1M4 repeat: matches 2983..3060 of consensus"
 716..852
 repeat_region
 /note=-52 copies 59 mer 95% conserved"
 1172..1289
 repeat_region
 /note=-2 copies 59 mer 96% conserved"
 1203..2147
 repeat_region
 /note=-180 copies 63 mer 93% conserved"
 1406..1800
 repeat_region
 /note=-match: STS: Em:H84161.187"
 2050..2167
 misc_feature
 /note=-2 copies 59 mer 96% conserved"
 1705..2167
 repeat_region
 /note=-2 copies 63 mer 95% conserved"
 1705..2167
 repeat_region
 /note=-L1M4 repeat: matches 2739..3049 of consensus"
 2673..3073
 repeat_region
 /note=-L1M4 repeat: matches 2..398 of consensus"
 3075..3755
 repeat_region
 /note=-L1M6 repeat: matches 2007..2681 of consensus"
 3757..4079
 repeat_region
 /note=-L1M6 repeat: matches 1..310 of consensus"
 4080..4199
 repeat_region
 /note=-L1M6 repeat: matches 1731..2007 of consensus"
 4702..5001
 repeat_region
 /note=-L1M6 repeat: matches 1..299 of consensus"
 5002..5199
 repeat_region
 /note=-L1M6 repeat: matches 1574..1731 of consensus"
 5633..5721
 repeat_region
 /note=-L1M6 repeat: matches 485..565 of consensus"
 6110..6232
 repeat_region
 /note=-L1M6 repeat: matches 2..132 of consensus"
 6961..7197
 repeat_region
 /note=-L12 repeat: matches 2265..2748 of consensus"
 7940..8022
 repeat_region
 /note=-MR repeat: matches 80..166 of consensus"
 8023..8074
 misc_feature
 /note=-"CG island"
 8075..8075
 gene
 /evidence=not_experimental
 join(8328..8770,8627..8733,36524..36951,37749..38394,
 39118..39225,42535..42684,46851..51301)
 mRNA
 join(8328..8770,8627..8733,36924..36951,37749..38394,
 39118..39225,42535..42684,46851..51301)
 /gene="b3314N13.5"
 /product="b3314N13.5 (synbaphillin (KIAA0374))"
 Em:AK004905
 Em:AK004905
 Em:AK002792
 match: ESTs: Em:AF1618.053 Em:AF207200 Em:U03740 Em:AI242746
 Em:H05310 Em:AB973638 Em:U15472 Em:AI3242482 Em:AI1221800
 Em:U133417 Em:K35903 Em:AI285452 Em:U04515 Em:AI759880"
 join(8627..8733,38298..38394)
 gene
 /gene="dJ54517.2"
 join(8627..8733,38298..38394)
 mRNA
 /product="b3314N13.1 (novel protein similar to RAD21 (S.
 pombe) homolog (KIAA0078))"
 /note=-match: ESTs: Em:AL042518 Em:AL049037"
 /evidence=not_experimental
 join(8627..8733,38298..38394)
 /gene="dJ54517.2.52366..38370"
 /note=-continues in Em:AL031665 as gene dJ54517.2 (gene
 fragment dJ54517.1)
 overlaps 5' end of gene .5"
 /evidence=not_experimental
 join(8627..8733,38298..38394)
 /product="b3314N13.1 (novel protein similar to RAD21 (S.
 pombe) homolog (KIAA0078))"
 /protein_id="CAC28880.1"
 /translation="MVGCGGAAAAATAPALRQCPGPGPCFAGLRASGLSPDPTATP
 "

Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- protect information -----

Center project name: G1YC
Center clone name: CH230-83A9

```

----- Summary statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye: 100% of reads

```

Assembly program: Phrap; version 0.990329
Consensus quality: 148489 bases at least Q
Consensus quality: 154533 bases at least Q

Consensus quality: 13455 bases at least Q30
Consensus quality: 160056 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
 (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)

NOTE: This is a 'working draft' sequence. It currently consists of 66 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will

* be preserved.
* 1 1233: contig of 1233 bp in length
* 1234

1234	1333: gap of unknown length
1334	2440: contig of 1107 bp in length
2441	2540: gap of unknown length

2441	gap of unknown length
2540	contig of 1042 bp in length
2541	contig of 1042 bp in length
3582	gap of unknown length
3583	gap of unknown length

*	3683	4827: contig of 1145 bp in length
*	4828	4927: gap of unknown length

* 4928 5942: contig of 1015 bp in length
* 5943 6042: gap of unknown length

```
* 6043 7144: contig of 1102 bp in length
* 7145 7244: gap of unknown length
```

* 7245 8917: contig of 1673 bp in length
* 8918 9017: gap of unknown length
* 8919 9017: gap of unknown length

* 9018	10051: contig of 1034 bp in length
* 10052	10151: gap of unknown length
* 10152	11346: contig of 1105 bp in length

- * 10152: contig of 1153 bp in length
- * 11346: gap of unknown length
- * 11347: contig of 1320 bp in length

12817	12916: gap of unknown length
12917	14061: contig of 1145 bp in length

*	14062	contig of 1900 bp in length
*	14161	gap of unknown length
*	14162	contig of 1900 bp in length

* 16062 16161: gap of unknown length
* 16162 17516: contig of 1355 bp in length

```

★ 17517 17616: gap of unknown length
★ 17617 19014: contig of 1398 bp in length

```

* 19015	19114: gap of unknown length
* 19115	20621: contig of 1507 bp in length

*	20622	20721: gap of unknown length
*	20722	21729: contig of 1008 bp in length
*	21730	21930: gap of unknown length

Accession	Gene	Length
21730	gap of unknown length	
21829	contig of 1506 bp in length	
21830	contig of 1506 bp in length	
23336	gap of unknown length	
23435	gap of unknown length	

23435	gap of unknown length
23436	contig of 1510 bp in length
24945	gap of unknown length
25045	gap of unknown length

* 25046 26552: contig of 1507 bp in length
* 26553 26652: gap of unknown length

* 26653 28252: contig of 1600 bp in length
* 28253 28352: gap of unknown length

* 28353 30791: contig of 2439 bp in length
* 30792 30891: gap of unknown length

33082	contig of 2197 bp in length
33083	gap of unknown length
33089	contig of 3015 bp in length
33203	gap of unknown length
33204	gap of unknown length
33624	contig of 1245 bp in length
33642	contig of 1245 bp in length
33643	contig of 1445 bp in length
40097	contig of unknown length
40181	gap of unknown length
40182	contig of 2034 bp in length
40183	gap of unknown length
42322	contig of 1951 bp in length
42323	contig of unknown length
43727	gap of unknown length
43728	contig of 2751 bp in length
43729	gap of unknown length
47124	gap of unknown length
47125	contig of unknown length
48727	gap of unknown length
48728	contig of 2675 bp in length
51402	contig of unknown length
51403	contig of 2196 bp in length
51502	gap of unknown length
51503	contig of 2116 bp in length
55914	gap of unknown length
55915	contig of unknown length
55916	contig of 1685 bp in length
57659	gap of unknown length
57660	gap of unknown length
60006	gap of unknown length
60007	contig of unknown length
60103	gap of unknown length
60104	contig of 2325 bp in length
63430	gap of unknown length
63431	contig of 2281 bp in length
63432	contig of 2226 bp in length
63433	contig of unknown length
63727	contig of 1959 bp in length
63728	gap of unknown length
63729	contig of 2547 bp in length
63730	gap of unknown length
71024	gap of unknown length
71025	contig of unknown length
71121	gap of unknown length
71122	contig of 2024 bp in length
71348	contig of unknown length
72349	gap of unknown length
72350	gap of unknown length
76023	gap of unknown length
76024	contig of 3411 bp in length
76464	gap of unknown length
76465	contig of unknown length
82146	contig of 2582 bp in length
82147	contig of 2547 bp in length
84793	contig of 2547 bp in length
84794	gap of unknown length
84879	contig of 2894 bp in length
84880	gap of unknown length
90047	gap of unknown length
90048	contig of 3601 bp in length
90249	contig of 3565 bp in length
90250	contig of 4535 bp in length
94785	gap of unknown length
94786	contig of 3501 bp in length
94884	contig of 3501 bp in length
94885	contig of 3501 bp in length
94886	contig of 3501 bp in length
102510	gap of unknown length
102511	contig of 3555 bp in length
103506	gap of unknown length
103507	contig of unknown length
103508	gap of unknown length
103509	contig of 3123 bp in length
112309	gap of unknown length
112310	contig of 3344 bp in length
112311	contig of 3344 bp in length
112312	contig of 3344 bp in length
112313	contig of 3314 bp in length

Query Match 4.3%; score 41; DB 2; Length 203466;
Best Local Similarity 47.2%; Pred. No. 1.6;

```
Matches 125; Conservative 0; Mismatches 140; Indels 0; Gaps 0;
```

QY 378 GGTCGAGTCGGGTCGGGTCAGGTATGATCGACCCGGTTATTTCTCCATGGGGTTTTGTT 437

Db 155176 GGTGGGGTCAGTTCTCCTACTCTCATGTCTCCAGAGCTACTCTCCCATGATGTCTAGGC 155235

[illegible]

* 107764 107663: gap of 100 bp
 * 145788: contig of 17825 bp in length
 * 145789 145788: contig of 10127 bp in length
 * 145789 180895: contig of 35107 bp in length.

FEATURES

Source

Location/Qualifiers

1..180895
 /organism="Mus musculus"
 /clone="t3non110050"

misc_feature

1..11367
 /clone_lib="RPC1-23 Female Mouse BAC"
 /note="assembly-fragment
 clone end:SP6"

misc_feature

11468..112983
 vector_side:left"

/note="assembly-fragment"

misc_feature

16486..47027
 /note="assembly-fragment"

misc_feature

47128..107763
 /note="assembly-fragment"

misc_feature

107884..145688
 /note="assembly-fragment"

misc_feature

145789..180895
 /note="assembly-fragment
 clone end"

misc_feature

vector_side:right"

BASE COUNT 54612 a 36242 c 35838 g 53603 t 600 others
 ORIGIN

Query Match 4.38; Score 40.2; Ds 2; Length 180895;
 Best Local Similarity 57.68; Pred.No. 2.6;
 Matches 72; Conservative 0; Mismatches 53; Indels 0; Gaps 0;
 Oy 115 CACTGACAGCTGCTGCGACGACGACGACGCTGCGAAGACGAGGAGAAC 174
 ||||| ||| ||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
 Db 24885 CACTGAGGAGAAATGTTAGAAAGAGGAGAGACAAAGGAGGAGGAGGAGGAG 24830
 Oy 175 CAGCTTCAGAGAGATGCTTACCAACCTCGACGCTCGAGAGATCGAAGAGAAAGA 224
 ||||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
 Db 24829 GAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGCA 24770
 Oy 235 AGCAA 239
 |||||
 Db 24765 AGAA 24765

Search completed: January 25, 2003, 23:45:15
 Job time : 4108 secs

Class: BAC ends High quality sequence stop: 592. Local/Qualifiers 1: 592		A BAC End Sequence Database for Identifying Minimal Overlaps in Arabidopsis Genomic Sequencing. Update 3 Unpublished (1997) Source: JOURNAL COMMENT	
/organism="Arabidopsis thaliana" /strain="Columbia" /db_xref="taxon:3702" /clone_lib="TAMU" /sex="hermaphrodite" /note="Vector: Belorachi; Site_1: HindIII; Site_2: HindIII ; Produced by Rod Ming, 176 t		The Institute for Genomic Research 9712 Medical Center Dr., Rockville, MD 20850, USA Tel: 301 938 0200 Email: rounseley@igrr.org Seq primer: M13 Reverse Clas: BAC ends High quality sequence stop: 508. Local/Qualifiers 1: 508	
FEATURES SOURCE		/organism="Arabidopsis thaliana" /strain="Columbia" /db_xref="taxon:3702" /clone_lib="TAMU" /sex="hermaphrodite" /note="Vector: Belorachi; Site_1: HindIII; Site_2: HindIII ; Produced by Rod Ming, 151 t	
BASE COUNT 195 a 127 c 3 others		BASE COUNT 158 a 117 c 82 g 151 t	
Query Match 33.18; Score 312.4; DB 17; Length 592; Best Local Similarity 82.18; Pred. No. 9.9e-92; Matches 412; Conservative 0; Mismatches 1; Indels 89; Gaps 1;		Query Match 31.48; Score 296.8; DB 17; Length 508; Best Local Similarity 82.04; Pred. No. 1.2e-91; Matches 408; Conservative 0; Mismatches 2; Indels 90; Gaps 2;	
Oy 529 AGAAGAACTTGGATGGATCAGATATAGTTCGATCAAGCGGTGGATTTT 588		Oy 531 AGAAACCTTGGATGGATCAGATATAGTTCGATCAAGCGGTGGATTTTCG 590	
Db 503 AGAAGAACTTGGATGGATCAGATATAGTTCGATCAAGCGGTGGATTTT 444		Db 508 AGAAGAACTTGGATGGATCAGATATAGTTCGATCAAGCGGTGGATTTTCG 449	
Oy 589 CGAATACAGATATCTCTCCGATGAAGCGGTAGTATCTTCTTCATCAG 648		Oy 591 AAATACATGATCTCTCTCGATGAAGCGGTAGTATCTTCTTCATCAGAT 650	
Db 443 CGAATACAGATATCTCTCCGATGAAGCGGTAGTATCTTCTTCATCAG 384		Db 446 ATATCATGATATCTCTCTCGATGAAGCGGTAGTATCTTCTTCATCAGAT 389	
Oy 645 ATCATCATGAGGACGAAGTTTCCTTATGATCATATATGCTAGACGATCAG 708		Oy 651 CATCATGAGGACGAAGTTTCCTTATGATCATATATGCTAGACGATCAGAT 710	
Db 383 ATCATCATGAGGACGAAGTTTCCTTATGATCATATATGCTAGACGATCAG 708		Db 386 CATCATGAGGACGAAGTTTCCTTATGATCATATATGCTAGACGATCAGAT 329	
Oy 709 TCTCTCTCTAGTACTATATCTCTTATGATCATATATGCTAGACGATCAG 763		Oy 711 TCTCTCTCTAGTACTATATCTTATGATCATATATGCTAGACGATCAGAT 763	
Db 323 TCTCTCTCTAGTACTATATCTTATGATCATATATGCTAGACGATCAG 264		Db 328 TCTCTCTCTAGTACTATATCTTATGATCATATATGCTAGACGATCAGAT 269	
Oy 764 ----- 763		Oy 764 ----- 763	
Db 263 TAAGTATGTCATTATTAATCACTATATATAGTATATATATATCTGTATCT 204		Db 266 AGTATGTCATTATTAATCACTATATATAGTATATATATATCTGTATCTTA 209	
Oy 764 ----- 763		Oy 764 ----- 763	
Db 203 TATTGATTAAGTCTGGGTTTAGGACCAATGAGGAAATGAGGAAATGAG 144		Db 208 TTTGATTAAGTCTGGGTTTAGGACCAATGAGGAAATGAGGAAATGAG 149	
Oy 800 AAACCTAGAAATGATCAGGAGGCTGCAAGGATACAGATTTTCCGGGAAATATGG 859		Oy 802 ACCCTAGAAATGATCAGGAGGCTGCAAGGATACAGATTTTCCGGGAAATATGG 860	
Db 143 AAACCTAGAAATGATCAGGAGGCTGCAAGGATACAGATTTTCCGGGAAATATGG 84		Db 148 ACCCTAGAAATGATCAGGAGGCTGCAAGGATACAGATTTTCCGGGAAATATGG 860	
Oy 860 TGAAGAGTTTCAAGTGGCTACACGCTGCTCACTGCTAGTATGCTATCAATAC 919		Oy 861 GAAGAGTTTCAAGTGGCTACACGCTGCTCACTGCTAGTATGCTATCAATAC 920	
Db 83 TGAAGAGTTTCAAGTGGCTACACGCTGCTCACTGCTAGTATGCTATCAATAC 24		Db 88 GAAGAGTTTCAAGTGGCTACACGCTGCTCACTGCTAGTATGCTATCAATAC 29	
Oy 920 CATGATATTCCTCTGAAGCT 941		Oy 921 ATTCATATTCCTCTGAAGCT 940	
Db 23 CATGATATTCCTCTGAAGCT 2		Db 28 ATTCATATTCCTCTGAAGCT 9	
RESULT 2 LOCUS B67977/1 508 bp DNA linear GSS 09-05C-1957 DEFINITION Arabidopsis thaliana genomic clone T35H20, DNA sequence. ACCESSION B67977 VERSION B67977.1 GI:2666731 SOURCE GSS ORGANISM Arabidopsis thaliana Eukaryota; Viridiplantae; Streptophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; rosales; Fabiales; Brassicales; Arabidopsi Rounsley S.D., Field C.E., Bass S., Linher K., Linher K., Golden K. AUTHORS Berry, K., Granger, D., Suh, E., Wible, C., Adams M. D. and Venter		RESULT 3 LOCUS B672044/3 541 bp DNA linear GSS 13 DEC-2001 DEFINITION Arabidopsis thaliana genomic clone B6H044, DNA sequence	

ESTION 4
SOURCE
ORGANISM

Brassica oleracea.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Posidae; Eurosida II; Brassicales; Brassicaceae; Eruicaceae;
1. (bases 1 to 641)
Whole genome shotgun sequencing of Brassica oleracea

Other-GSSs: B08H437F
Contact: Chris Town
7712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: ctown@nsl.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TR
Class: sheared ends.

REFERENCE
TICR
7712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: ctown@nsl.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TR
Class: sheared ends.

FEATURES
source
1. 775
/organism="Brassica oleracea"
/strain="T01000D93"
/clone="B08H437F"
/note="Vector: pBIO1; Site 1: BstXI; 2-3 kb sheared genomic DNA inserted into pBIO1 using BstXI linkers"

BASE COUNT
155 a 169 c 140 g 177 t

Query Match
Best Local Similarity 19.3%; Score 182.2; DB 17; Length 641;
Matches 367; Conservative 68.6%; Pred. No. 6.5e-43;
Y 46 GAAACCAACAGCTCTTCGAGAACACCGCTGTGTGCTAAGCTCCGCGAGATCC 105
b 641 GAAACCAACAGCTCTTCGAGAACACCTACCAAGCTCTTCAA-----CGCGAGATCC 588
Y 106 GCGCAGAGACACTCA--AGAGTCTGTGTCGAGAACAGCTGAGACAGCTGACGAAA 162
b 587 GCGCAGAGACACTCA--AGAGTCTGTGTCGAGAACAGCTGAGACAGCTGACGAAA 528
Y 163 AGAGAGAGAGACAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 222
b 527 AGAGAGAGAGACAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 468
Y 223 AGAGAGAGAGACAGCTGCGCGCGCTCACTCGAGACAGCTCACTAGCATCA 282
b 467 AG 413
Y 283 TCTCTACACAGCTACCGGTTTACCGGTTACCGGTTACCGGTTACCGGTTACCGGTT 342
b 412 -GAGCGCTACAGTCAAGCTGTACCGGTTACCGGTTACCGGTTACCGGTTACCGGTT 394
Y 343 TCCCAAGCTCACTGCG-----GGAGCAACA 366
b 353 TTCACAGCTACGGTGTGGAGCAACACAGCTCTCTCGAGGGGTATCTACTGAGAGCT 294
Y 367 GATCATATGTGTGGAGTGGGTGGGTGAGTATGATGAGTGGGTATGATGAGTGGGT 426
b 293 TCGTTCGCGGGAGAGAGCGGCTGGGTGAGATCATGATGAGAGCGGCTGTCTCTCT 234
Y 427 GGGGTTTGTGGAGCT 480
b 233 GGGGTTTGTGGAGATCT 186
Y 487 TGTTCATGCT 541
b 175 TGTATAGG---CTTCCATATATCTGAGATCTCTCTCTCTCTCTCTCTCTCTCTCT 128

RESULT 4

BH472100/c
DEFINITION

Brassica oleracea.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Posidae; Eurosida II; Brassicales; Brassicaceae; Eruicaceae;
1. (bases 1 to 775)
Whole genome shotgun sequencing of Brassica oleracea

Other-GSSs: B08H437F
Contact: Chris Town
7712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: ctown@nsl.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TR
Class: sheared ends.

REFERENCE
TICR
7712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: ctown@nsl.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TR
Class: sheared ends.

FEATURES
source
1. 775
/organism="Brassica oleracea"
/strain="T01000D93"
/clone="B08H437F"
/note="Vector: pBIO1; Site 1: BstXI; 2-3 kb sheared genomic DNA inserted into pBIO1 using BstXI linkers"

BASE COUNT
265 a 139 c 135 g 238 t

Query Match
Best Local Similarity 6.3%; Score 59.4; DB 17; Length 775;
Matches 115; Conservative 0; Mismatches 81; Indels 9; Gaps 1;

Y 760 ATACGAGCAATGAGGAGGATTTGGAGCTACATGAGAGAAACCTGAGATGATGAG 819
b 349 ATAGGCTCTCATGAGAGCTCTGGAGC-----GAGAGCTTACAGAGCAACA 299
Y 820 GAGGTGTGAGAGTACAGCTTTTTCGGGGAAATATGTAAGAGTTTTCAGTGTG 879
b 298 GAGGTGTGAGAGTACAGCTTTTTCGGGGAAATATGTAAGAGTTTTCAGTGTG 239
Y 880 CTACAAGCTGCTACTGTTGATGAGTGGAGTCTATGATGATGATGATGATGATGAT 939
b 238 TGTCTACATCAAGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 179
Y 940 TTTAA 944
b 178 TTTAA 174

RESULT 5
LOCUS
DEFINITION

Brassica oleracea.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Posidae; Eurosida II; Brassicales; Brassicaceae; Eruicaceae;
1. (bases 1 to 645)
Whole genome shotgun sequencing of Brassica oleracea

Other-GSSs: B08H437F
Contact: Chris Town
7712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: ctown@nsl.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TR
Class: sheared ends.

REFERENCE
TICR
7712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: ctown@nsl.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TR
Class: sheared ends.

FEATURES
source
1. 645
/organism="Brassica oleracea"
/strain="T01000D93"
/clone="B08H437F"
/note="Vector: pBIO1; Site 1: BstXI; 2-3 kb sheared genomic DNA inserted into pBIO1 using BstXI linkers"

BASE COUNT
265 a 139 c 135 g 238 t

Query Match
Best Local Similarity 6.3%; Score 59.4; DB 17; Length 775;
Matches 115; Conservative 0; Mismatches 81; Indels 9; Gaps 1;

Y 760 ATACGAGCAATGAGGAGGATTTGGAGCTACATGAGAGAAACCTGAGATGATGAG 819
b 349 ATAGGCTCTCATGAGAGCTCTGGAGC-----GAGAGCTTACAGAGCAACA 299
Y 820 GAGGTGTGAGAGTACAGCTTTTTCGGGGAAATATGTAAGAGTTTTCAGTGTG 879
b 298 GAGGTGTGAGAGTACAGCTTTTTCGGGGAAATATGTAAGAGTTTTCAGTGTG 239
Y 880 CTACAAGCTGCTACTGTTGATGAGTGGAGTCTATGATGATGATGATGATGATGAT 939
b 238 TGTCTACATCAAGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 179
Y 940 TTTAA 944
b 178 TTTAA 174

RESULT 5
LOCUS
DEFINITION

Brassica oleracea.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Posidae; Eurosida II; Brassicales; Brassicaceae; Eruicaceae;
1. (bases 1 to 645)
Whole genome shotgun sequencing of Brassica oleracea

Query Match 4.1% Score 39; DB 14; Length 519;
 Best Local Similarity 54.3%;
 Matches 78; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

Y 548 TGATCAGATATATGTTAGTTCGATCAGCGGTGGTGGATTTTCGAATACATCATGATTC 607
 b 359 TATACATATTATCCACACACAGTGGTGGTGTATATATATTCACAAATATACC 300
 Y 608 TCTCCGATCAGCGGTGATCAGTATCTTCTTCATCAGATCATCATCAGAGACCA 667
 b 299 TCCACCAATATAGCTATATAGTATGTTAGTTCGCCACCTCTGTATCCATATACC 240
 Y 668 AGGTTCTCTTATGATCATAGAA 690
 b 239 TTGGTGGCTTATCATATAGAA 217

RESULT 12
 1432782
 DEFINITION
 PPRV3543 P. infestans-challenged leaf Solanum tuberosum cDNA clone
 PPRV3543 5' sequence, mRNA sequence.
 ACCESSION
 EMISSION
 PROS
 KEYWORDS
 ORGANISM
 Solanum tuberosum
 Bacteria; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Asteridae; eustersids I; Solanales; Solanales; Solanaceae; Solanum.
 1 (bases 1 to 522)
 AUTHORS
 Bougri, O., Buell, C.R., Renning, C.M., Fry, M.E. and Baker, B.
 TITLE
 Infestans-challenged leaves Challenged with Phytophthora
 Infestans Compatible Interaction
 JOURNAL
 Unpublished (2000)
 COMMENT
 Contact: Cathy Renning
 The Institute for Genomic Research
 Division tel 1-800-711-6195, email cdraegreen.com
 Seq primer: M13P-R.

FEATURES
 source
 Location/Qualifiers
 1..522
 /organism="Solanum tuberosum"
 /cultivar="Kennebec"
 /db_xref="taxon:4113"
 /clone="PPRV35"
 /clone.lib="infestans-challenged leaf"
 /dev_stage="6 week old"
 /lab_host="SOLR"
 Note-Vector: pBluescript SK(-); Site.1: EcoRI; Site.2:
 Approx.1500 bp. Infestans US-105 940501 in Blotter
 (Madison, Wisconsin). Leaf tissue was collected at 1, 2,
 5, 12, and 24 hours post-challenge and frozen in liquid
 nitrogen immediately upon removal. Kennebec plants showed
 no symptoms of infection. Infestans US-1 were used as controls and showed
 infection. NOTE: We cannot exclude the possibility that
 this sequence is actually derived from Phytophthora rather
 than potato. 63 g 192 t

BASE COUNT
 BASE COUNT 139 a 126 c 63 g 192 t

Query Match 4.1% Score 39; DB 13; Length 522;
 Best Local Similarity 54.3%;
 Matches 78; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

Y 548 TGATCAGATATGTTAGTTCGATCAGCGGTGGTGGATTTTCGAATACATCATGATTC 607
 b 161 TATACATATTATCCACACACAGTGGTGGTGTATATATATTCACAAATATACC 220

Y 608 TGCTCCGATCAGCGGTGATCAGTATCTTCTTCATCAGATCATCATCAGAGACCA 667
 b 221 TCCACCAATATAGCTATATAGTATGTTAGTTCGCCACCTCTGTATCCATATACC 280
 Y 668 AGGTTCTCTTATGATCATAGAA 690
 b 281 TTGGTGGCTTATCATATAGAA 303

RESULT 13
 1432782
 DEFINITION
 PPRV3540 P. infestans-challenged leaf Solanum tuberosum cDNA clone STM1550
 PPRV3540 5' sequence, mRNA sequence.
 ACCESSION
 EMISSION
 PROS
 KEYWORDS
 ORGANISM
 Solanum tuberosum
 Bacteria; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Asteridae; eustersids I; Solanales; Solanales; Solanaceae; Solanum.
 1 (bases 1 to 525)
 AUTHORS
 Restrepo, S., Griffiths, H., van der Hoeven, R., Tsai, J. and
 Karanycheva, S.A.
 TITLE
 Generation of a set of potato cDNA clones for microarray analyses
 Unpublished (2002)756
 COMMENT
 Contact: Robin Buell
 The Institute for Genomic Research
 9712 Medical Center Dr, Rockville, MD 20850, USA
 This clone is available through the Research Genetics, contact the
 Research Genetics for further information 1-800-711-6195 or
 cdraegreen.com
 Seq primer: T1.

FEATURES
 source
 Location/Qualifiers
 1..525
 /organism="Solanum tuberosum"
 /cultivar="Kennebec or Bintje"
 /db_xref="taxon:4113"
 /clone.lib="STM1550"
 /clone.lib="Generation of a set of potato cDNA clones for
 microarray analyses mixed potato tissues"
 /tissue_type="mixed tissues"
 /lab_host="SOLR"
 Note-Vector: pBluescript SK(-); Site.1: EcoRI; Site.2:
 XhoI; supplier: Combination of untreated and Phytophthora
 infestans-treated libraries of stolons, leaves, leaflets,
 axillary buds of stem explants, petioles, germinating eyes
 of tubers, or roots.
 165 a 120 c 99 g 141 t

BASE COUNT
 BASE COUNT 165 a 120 c 99 g 141 t

Query Match 4.1% Score 39; DB 14; Length 525;
 Best Local Similarity 54.3%;
 Matches 78; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

Y 548 TGATCAGATATGTTAGTTCGATCAGCGGTGGTGGATTTTCGAATACATCATGATTC 607
 b 251 TATACATATTATCCACACACAGTGGTGGTGTATATATTCACAAATATACC 310
 Y 608 TCTCCGATCAGCGGTGATCAGTATCTTCTTCATCAGATCATCATCAGAGACCA 667
 b 311 TCCACCAATATAGCTATATAGTATGTTAGTTCGCCACCTCTGTATCCATATACC 370
 Y 668 AGGTTCTCTTATGATCATAGAA 690
 b 371 TTGGTGGCTTATCATATAGAA 393

2000

This Page Blank (uspio)

Matches 314; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATSLFPMSTQNSGNPNDLNTLRYVNSGSEIFETLASKRGKSGKTKGQOKKEPT 60

DY 1 MATSLFPMSTQNSGNPNDLNTLRYVNSGSEIFETLASKRGKSGKTKGQOKKEPT 60

DY 61 LAGWAKLEPQRIEERKQALAAATGDTSSVASSNATRLPVPDGVYLGQFPSSLG 120

DY 61 LAGWAKLEPQRIEERKQALAAATGDTSSVASSNATRLPVPDGVYLGQFPSSLG 120

DY 61 LAGWAKLEPQRIEERKQALAAATGDTSSVASSNATRLPVPDGVYLGQFPSSLG 120

DY 121 SRTYCGSGQWMDPVSFPGVETSSSTHLESIINPQFNASSNNKDCCTFKKR 180

DY 121 SRTYCGSGQWMDPVSFPGVETSSSTHLESIINPQFNASSNNKDCCTFKKR 180

DY 181 LQDQNNVNSGGSGFKYTMPPPNKGDTQYLQSDHHSOGFLYDRIARAASVSAS 240

DY 181 LQDQNNVNSGGSGFKYTMPPPNKGDTQYLQSDHHSOGFLYDRIARAASVSAS 240

QY 241 STTNFYFNATNTHGEMFSGYTHGPNRSGSGVYEFYFGYGRVSVATSSLV 300

DY 241 STTNFYFNATNTHGEMFSGYTHGPNRSGSGVYEFYFGYGRVSVATSSLV 300

QY 301 GDCSPNTIDLSKL 314

DY 301 GDCSPNTIDLSKL 314

DY 301 GDCSPNTIDLSKL 314

RESULT 2

081836 PRELIMINARY: PRT: 314 AA.

ID 081836

AC 01-NOV-1998 (TREMREL. 08, Created)

DT 01-NOV-1998 (TREMREL. 08, Last sequence update)

DT 01-OCT-2000 (TREMREL. 15, Last annotation update)

DT 01-OCT-2000 (TREMREL. 15, Last annotation update)

DT 01-OCT-2000 (TREMREL. 15, Last annotation update)

DT 01-OCT-2000 (TREMREL. 15, Last annotation update)

DT 01-OCT-2000 (TREMREL. 15, Last annotation update)

DT 01-OCT-2000 (TREMREL. 15, Last annotation update)

DT 01-OCT-2000 (TREMREL. 15, Last annotation update)

DT 01-OCT-2000 (TREMREL. 15, Last annotation update)

DT 01-OCT-2000 (TREMREL. 15, Last annotation update)

DT 01-OCT-2000 (TREMREL. 15, Last annotation update)

DT 01-OCT-2000 (TREMREL. 15, Last annotation update)

DT 01-OCT-2000 (TREMREL. 15, Last annotation update)

DT 01-OCT-2000 (TREMREL. 15, Last annotation update)

DT 01-OCT-2000 (TREMREL. 15, Last annotation update)

DT 01-OCT-2000 (TREMREL. 15, Last annotation update)

DT 01-OCT-2000 (TREMREL. 15, Last annotation update)

DT 01-OCT-2000 (TREMREL. 15, Last annotation update)

DT 01-OCT-2000 (TREMREL. 15, Last annotation update)

DT 01-OCT-2000 (TREMREL. 15, Last annotation update)

DT 01-OCT-2000 (TREMREL. 15, Last annotation update)

DT 01-OCT-2000 (TREMREL. 15, Last annotation update)

DT 01-OCT-2000 (TREMREL. 15, Last annotation update)

DT 01-OCT-2000 (TREMREL. 15, Last annotation update)

DT 01-OCT-2000 (TREMREL. 15, Last annotation update)

DT 01-OCT-2000 (TREMREL. 15, Last annotation update)

DT 01-OCT-2000 (TREMREL. 15, Last annotation update)

DT 01-OCT-2000 (TREMREL. 15, Last annotation update)

DT 01-OCT-2000 (TREMREL. 15, Last annotation update)

DT 01-OCT-2000 (TREMREL. 15, Last annotation update)

DT 01-OCT-2000 (TREMREL. 15, Last annotation update)

DT 01-OCT-2000 (TREMREL. 15, Last annotation update)

DT 01-OCT-2000 (TREMREL. 15, Last annotation update)

[illegible][illegible]

[illegible][illegible]

```

Spir E., Spradling A.C., Stapleton M., Strong R., Sun E.,
Svitskas R., Tector C., Turner B., Venter B., Wang A.H., Wang X.,
Wang Z.-Y., Wattarun D.A., Weinstock G.M., Wessienbach J.,
Williams S.M., Woodage I., Worley K.C., Wu D., Yang S., Yao C.A.,
Zhang H.F., Zhang Q., Zhu S., Zhou X., Zhu S., Zhu X., Smith H.O.,
Zheng X.H., Zheng F.N., Zhong W., Zhou X., Zou J.C.:
Gibbs R.A., Myers E.R., Rubin G.M., Venter J.C.:
"The genome sequence of Drosophila melanogaster";
Science 287(5414):15-16 (2000).
Sequence ID: 15458.5; 2193(2000).
Accession: U00000.1; 15458.5.
FlyBase: FBgn020930; cngl.
InterPro: IPR000593; cNMP-binding.
InterPro: IPR001622; K-channel_iPore.
InterPro: IPR000638; M-channel_iPore.
InterPro: IPR002422; Voltage-gated_K_channel_1.
Pfam: PF00320; Ion_Trans_1.
SMART: SM00100; cNMP: 1.
PROSITE: PS00888; cNMP_BINDING: 1.
PROSITE: PS00889; cNMP_BINDING_2: 1.
PROSITE: PS50042; cNMP_BINDING_3: 1.
SEQUENCE 1693 AA: 182854 MW: 637857836975CDA CRC64:
Query Match      6.34; Score 103; DB 5; Length 1693;
Query Local Similarity 12.44; Pred. MO: 2;
Matches 7; 44; Mismatches 131; Indels 108; Gaps
15
7  FNSTDQNSVGNPNLRLTVLVSSGKATLTKEKRKK-----PSGKTQQQKO 56
964  LLETTTSTTS--TGTTSTSTASASASASASNGRQGYVPDLPSPENETWKRR 1022
57  KYFTLGAGVKALIPROITEPKQLAAATGVDSVASINATPLVPVDVPGVLGFP 116
1023  LXILIKLLMGQGGQEGAGTGTGGAGMYDTGSKNVSARSKO CIAPAP-TLEGPP 1079
117  SS-----LSGRNYSI-GVGSGQWDIVDPISPMGFVEISTHSELSSINQ 162
1080  ASAMEEESFGGAAAGSVNNNAACNCMSGV-----CNGCSNTVNDESA ----- 1126
163  WFHASNWRNDVCYFKKLDDQNVNVSNGSGSKYFMDPKMWTQYLLOSHDWKS 222
1127  -HNSSH-----PSNN-----GNSATCTPEP-----S 1147
223  GATFDYERIAAAS-----YVASSTTINPYFNMXITNIGMEPF-----GSYME 266
1148  VFPEPVPIAPASSGVPLPGLVSSNTPSLVTITNTPSQSFAPALQELGKHGT 1207
267  ONFRNGSGCKVYFFPKTKGERVSNVATTS--SLVGDKS-----PTHTIDSLK 314
1208  GRULGVGVGASNSNMKAGASASASASTTANYSIVFLCSASNSNIHLNHSHTI 1267

```

Q870H7	PRELIMINARY:	PRT:	406 AA.
Q870H7	0002 (T-EMBLrel 21, Created)		
	01-JUN-2002 (T-EMBLrel 21, Last annotation update)		
	01-JUN-2002 (T-EMBLrel 21, Last annotation update)		
	LD023499.		
	CSP.		
	Asaphidella malinogaster (Fruit fly)		
	Phlebotominae: Phlebotominae		
	Pterygota: Mezoptera: Arthropoda: Tracheata: Hexapoda: Insecta:		
	Pterygota: Neoptera: Endopterygota: Diptera: Brachycera: Muscomorpha:		
	Ephydroidea: Drosophilidae: Drosophila.		
	NCBI_TaxID=7227:		
	SEQUENCE FROM A		
	STRAIN=RCFLEY:		
	Stapleton M., Brokstein P., Hong L., Agbanyani A., Carlson J.,		
	Champe M., Chavez C., Dorsett V., Fartán D., Frise E., George R.,		
	Gonszler M., Guarin H., Hill P., Liao G., Miranda A., Mungall C.J.,		
	Nguyen T., O'Connell M., O'Connell M., O'Connell M., O'Connell M.,		
	Yu C., Lewis S.E., Robin G.M., Cepliker S.		

[illegible]

RESULT 5

T	variation	replace(169054,G)	*tag= a	
T		/standard_name= "Single nucleotide polymorphism"		
T	intron	169121..169542		
T		/number= 5		
T	exon	/cons_splice= (5'site:YES,3'site:NO)		
T		169543..169925		
T		*tag= b		
T	variation	replace(169912,G)		
T		*tag= a		
T	variation	/standard_name= "Single nucleotide polymorphism"		
T		replace(169922,G)		
T		*tag= b		
T	intron	/standard_name= "Single nucleotide polymorphism"		
T		169926..169986		
T		*tag= b		
T	exon	/cons_splice= (5'site:YES,3'site:NO)		
T		169987..170678		
T		*tag= b		
T	variation	/number= 170679		
T		/number= (170035..A)		
T		*tag= a		
T	variation	/standard_name= "Single nucleotide polymorphism"		
T		replace(170066,T)		
T	variation	/standard_name= "Single nucleotide polymorphism"		
T		replace(170255,C)		
T		*tag= a		
T	variation	/standard_name= "Single nucleotide polymorphism"		
T		replace(170366,G)		
T		*tag= a		
T	variation	/standard_name= "Single nucleotide polymorphism"		
T		replace(170487,C)		
T		*tag= a		
T	intron	/standard_name= "Single nucleotide polymorphism"		
T		170679..204911		
T		*tag= b		
T	exon	/number= 7		
T		204912..205102		
T		*tag= b		
T	intron	/number= 8		
T		205103..242969		
T		*tag= b		
T	exon	/number= 9		
T		242970..243086		
T		*tag= b		
T	variation	replace(243055,T)		
T		*tag= a		
T	intron	/standard_name= "Single nucleotide polymorphism"		
T		243087..306167		
T		*tag= b		
T	variation	/number= 9		
T		replace(243187,C)		
T		*tag= a		
T	exon	/standard_name= "Single nucleotide polymorphism"		
T		306168..306503		
T		*tag= b		
T	variation	/number= 10		
T		replace(306592,A)		
T		*tag= a		
T	variation	/standard_name= "Single nucleotide polymorphism"		
T		replace(306383,G)		
T		*tag= a		
T	intron	/standard_name= "Single nucleotide polymorphism"		
T		306504..373639		
T		*tag= b		
T	exon	/number= 10		
T		373640..373778		
T		*tag= b		

PT			/number= 11	
PT	Intron		373779..422963	
PT			/tag= b	
PT	exon		/number= 11	
PT			422964..423097	
PT			/tag= a	
PT			/number= 12	
PT	variation		replace(423067..C)	
PT			/tag= a	
PT			/standard_name= "Single nucleotide polymorphism"	
PT	Intron		423098..459535	
PT			/tag= b	
PT			/number= 12	
PT	variation		replace(423149..G)	
PT			/tag= a	
PT			/standard_name= "Single nucleotide polymorphism"	
PT	variation		replace(423193..G)	
PT			/tag= a	
PT			/standard_name= "Single nucleotide polymorphism"	
PT	variation		replace(423220..A)	
PT			/tag= a	
PT			/standard_name= "Single nucleotide polymorphism"	
PT	variation		replace(423232..G)	
PT			/tag= a	
PT			/standard_name= "Single nucleotide polymorphism"	
PT	variation		replace(423258..G)	
PT			/tag= a	
PT			/standard_name= "Single nucleotide polymorphism"	
PT	exon		456354..456537	
PT			/tag= b	
PT			/number= 13	
PT	Intron		456538..460700	
PT			/tag= a	
PT			/number= 13	
PT	variation		replace(459706..C)	
PT			/tag= a	
PT			/standard_name= "Single nucleotide polymorphism"	
PT	variation		replace(459832..A)	
PT			/tag= a	
PT			/standard_name= "Single nucleotide polymorphism"	
PT	variation		replace(459913..G)	
PT			/tag= a	
PT			/standard_name= "Single nucleotide polymorphism"	
PT	variation		replace(460024..G)	
PT			/tag= a	
PT			/standard_name= "Single nucleotide polymorphism"	
PT	variation		replace(460056..T)	
PT			/tag= a	
PT			/standard_name= "Single nucleotide polymorphism"	
PT	variation		replace(460199..G)	
PT			/tag= a	
PT			/standard_name= "Single nucleotide polymorphism"	
PT	variation		replace(460553..C)	
PT			/tag= a	
PT			/standard_name= "Single nucleotide polymorphism"	
PT	variation		replace(460584..A)	
PT			/tag= a	
PT			/standard_name= "Single nucleotide polymorphism"	

[illegible]

Y 163 AACAGAGAACAACAGCTTCAGAGAGAAATGGCTGTAGCAAGCTCGAGGCTCAGAGAAATCG 222
 b 310 GAAG 251
 Y 223 AAG 239
 b 250 AAG 234
 RESULT 9
 BO70984
 C ABO70984 standard; DNA; 4604 BP.
 X ABO70984;
 X
 Y 29-AUG-2002 (first entry)
 X
 b Listeria monocytogenes 4b contig DNA sequence #926.
 Y Anthracis; Listeria; food contamination; mutational analysis;
 W infection; ds.
 X
 Y Listeria monocytogenes 4b.
 X
 N WO200228891-A2.
 X
 P 11-APR-2002.
 X
 F 04-OCT-2001; 2001WO-PR03061.
 X
 R 04-OCT-2000; 2000PR-0012697.
 X
 A (INSP) INST PASTEUR.
 A (CNRS) CNRS CENT NAT RECH SCI.
 X Kunst F, Glaser P;
 X
 R WPI: 2002-312479/37.
 X
 T New genomic sequences from *Listeria* species, useful for detection,
 T diagnosis and epidemiology of infection; also related polypeptides,
 T antibodies and modulators
 X
 S Claim 14; SEQ ID 3797; 180pp; French.
 X
 C The present invention relates to nucleic acid sequences
 C (AB057188-AB071212) from *Listeria* sp. The sequences are useful as probes
 C and primers for identification and/or detection of *Listeria* (e.g. as
 C contaminants in foods, or mutational analysis) and for analysis of
 C gene expression. Proteins encoded by the nucleic acid sequences can be
 C used for diagnosis and epidemiology of infection, for immunization
 C and pathogenicity of *Listeria* (potential therapeutic agents), also for
 C treating infections by *Listeria*, and are useful as immunogens in
 C anti-*Listeria* vaccines.
 C Note: The sequence data for this patent did not form part
 C of the prior art.
 C directly from NIPO at http://wipo.int/pub/published_pct_sequences.
 X
 Q Sequence 4604 BP; 1433 A; 774 C; 1119 G; 1278 T; 0 other:
 Query Watch 3.7%; Score 34.6; DB 24; Length 4604;
 Best Local Similarity 52.4%; Pred. No. 3; DB
 Matches 99; Conservative 0; Mismatches 89; Indels 1; Gaps 1;
 Y 751 GACACATATACAGGAGACATACAGAGATTTGGACATCTAGTCAGAGAAACCTAGAA 810
 b 4051 CAAAAACGTTTAAAGCAAGCTGACGATTTGGCGACAGTATGACGAGAGAGCCCTTGAG 4110
 Y 811 ATGATACAGAGGTGTCGAAGAGATACAGATTTTTTCGGGGGAAATATGTCGAAGATTTT 870
 b 4111 ATGCTTGAAATGAAGAGCGAGGATCTATATATGCTTTTGTATGGGCAAGAGATTT 4170

OY 871 CAGTCGTGGCTACAGCTGCTGCTAGCTAGGCTGATCGAGTCTAATACCATGATTCT 930
 Db 4171 CGAGTTTCACTCCACTTCATCAGGCTGCGCATTCGA-TCTTTTACCATGATTATTCT 4229
 OY 931 CTTCTGACGC 939
 Db 4230 CGATGTACC 4238
 RESULT 10
 AAQ95196
 ID AAQ95196 standard; DNA; 494 BP.
 X AAQ95196;
 X
 DT 08-FEB-1996 (first entry)
 XX
 XX Simple tandem repeat (STR) corresponding to wg15.
 XX Simple tandem repeat; STR; wg15; treatment; genetic; diagnosis;
 KW characterisation; mapping; linkage studies; analysis; alleles; ss.
 XX Synthetic.
 XX
 PN WO9517522-A2.
 XX
 XX 29-JUN-1995.
 PT 21-DEC-1994; 94WO-GB02789.
 XX
 PR 21-DEC-1993; 93CP-0026052.
 XX
 XX (UYLE-) UNIV LEICESTER.
 XX
 PT Armour J, Jeffreys AJ;
 XX WPI: 1995-240682/31.
 PT Identifying simple tandem repeat loci in DNA - by screening DNA
 PT library to enrich for fragments contg the repeats before cloning
 PT and rescreening, also simple tandem repeats for treatment or
 PT diagnosis
 XX
 PS Claim 26; Pages 23-24; 51pp; English.
 CC AAQ95196 is a simple tandem repeat (STR) corresponding to wg15. The
 CC sequence is useful for genetic characterisation, mapping, linkage
 CC medicine, partic for genetic characterisation, mapping, linkage
 CC studies and analysis/diagnosis of acquired disease alleles.
 SO Sequence 494 BP; 223 A; 72 C; 140 G; 44 T; 15 other:
 Query Watch 3.6%; Score 34.2; DB 16; Length 494;
 Best Local Similarity 49.0%; Pred. No. 1.6;
 Matches 72; Conservative 0; Mismatches 75; Indels 0; Gaps 0;
 OY 99 GAGATCTCGGACAGACACTGTAGAGCTGCTGCGAAACGAGGTCGAGAGAGCTCAG 156
 Db 156 GAAACACGACAGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 215
 OY 159 CAAAACAGAGAGAAACGCTGTGAGAGCAATGGCTGTGCAAACTCGAGCTCAGAG 218
 Db 216 GAGGACGA 275
 QY 219 ATCGAAGAGAGAAAGAACCACTGGCC 245
 Db 276 ACGAATATNNNNNNNNNNNACTCC 302
 RESULT 11
 AAS16356
 ID AAS16356 standard; cDNA; 1292 BP.
 X

Query Match 3.6%; Score 34.2; DB 22; Length 4221;
 Match Local Similarity 52.4%; Pred. No. 4.9;
 Bits 75; Conservative 0; Mismatches 68; Indels 0; Gaps 0;
 y 730 AGTACTACTATCTATCTTATTCACAGGACCAATCAATGAGGACCAATGAGGAA 779
 y 780 ATTCCTATCTATCTATCTTATTCACAGGACCAATCAATGAGGACCAATGAGGAA 829
 y 780 ATTCCTATCTATCTATCTTATTCACAGGACCAATCAATGAGGACCAATGAGGAA 879
 b 2140 TATTGACCAACACACATGATCCAGATATGAGGACCAACCAACGATGAG 2199
 y 840 TTTTTCGGGGAATGTGCA 862
 b 2200 TCTTTCCAGAAACTGTGCA 2222

RESULT 13
 D AAK80081
 X AAK80081 standard: DNA; 11363 BP.
 X AAK80081
 T 07-NOV-2001 (first entry)
 X Human immune/haematopoietic antigen genomic sequence S80 ID NO:34893.
 E Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
 W cytostatic; gene therapy; vaccine; metastasis; ds.
 X Homo sapiens.
 X WO200117182-A2.
 X 09-AUG-2001.
 X 17-JAN-2001; 2001WO-US01354.
 X 31-JAN-2000; 2000US-0175065.
 X 04-FEB-2000; 2000US-0180628.
 X 02-MAR-2000; 2000US-0186550.
 X 16-MAR-2000; 2000US-0189874.
 X 17-MAR-2000; 2000US-0190076.
 X 18-APR-2000; 2000US-0198123.
 X 07-JUN-2000; 2000US-0209467.
 X 28-JUN-2000; 2000US-0214886.
 X 30-JUN-2000; 2000US-0215135.
 X 07-JUL-2000; 2000US-0216647.
 X 11-JUL-2000; 2000US-0217487.
 X 14-JUL-2000; 2000US-0217496.
 X 14-JUL-2000; 2000US-0218290.
 X 26-JUL-2000; 2000US-0220063.
 X 14-AUG-2000; 2000US-0224518.
 X 14-AUG-2000; 2000US-0224519.
 X 14-AUG-2000; 2000US-0225213.
 X 14-AUG-2000; 2000US-0225266.
 X 14-AUG-2000; 2000US-0225267.
 X 14-AUG-2000; 2000US-0225268.
 X 14-AUG-2000; 2000US-0225270.
 X 14-AUG-2000; 2000US-0225271.
 X 14-AUG-2000; 2000US-0225277.
 X 14-AUG-2000; 2000US-0225278.

[illegible]

Search completed: January 25, 2003, 22:59:40
 Job time : 1723 secs

imis Page Blank (uspio)

Best Local Similarity 8.5%, Pred. No. 9.3e-05:

Matches 17; Conservative 68; Mismatches 68; Indels 0; Gaps 0:

QY 96 GCGAGATCGGACAGACACTAGAGACTGCTGCGGAACACAGCAAGCAAGCT 155
 Db 1236 RR 1177
 QY 156 CAGCAAAACAGAGAACCAACCTGTGAGCAAACTGGCTAGCAAGCTCGAGGTCAG 215
 Db 1176 RR 1117
 QY 216 AGATGAGAGAGAAAGACAGACTCGCGCCGACAGTGGAGACAGCTCATCAGTA 275
 Db 1116 RR 1111
 QY 276 GCATGATCTCTCAACAGCT 296
 Db 1056 CCTTCAGCTCGACCAAGCT 1036

RESULT 2

US-08-232-463-14
 Sequence 14 Application US/08232463
 Sequence 146047
 GENERAL INFORMATION:
 APPLICANT: DORNER, F.
 APPLICANT: SCHEIFELINGER, F.
 TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
 NUMBER OF SEQUENCES: 52
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Wiley & Lardner
 ADDRESS: 1800 Personal Road, Suite 500
 CITY: Alexandria
 STATE: VA
 COUNTRY: USA
 ZIP: 22304
 COMPUTER READABLE FORM:
 HDION TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 APPLICATION DATE: 01-08-1994
 CURRENT APPLICATION NUMBER: US/08/232.463
 FILING DATE: 01-08-1994
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/07/935.313
 FILING DATE: 01-08-1993
 APPLICATION NUMBER: EP 91 114 300.6
 FILING DATE: 01-08-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: BENT, Stephen A.
 REGISTRATION NUMBER: 29,768
 REFERENCE/DOCKET NUMBER: 30472/114 TMU
 TELEPHONE: (703) 836-5300
 TELEFAX: (703) 683-4109
 TELEX: 899149
 INFORMATION FOR SEQ ID NO: 14:
 SEQUENCE LENGTH: 7218 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 IMMUNOLOGY SOURCE: near
 CLONE: pT98pt-Fls
 US-08-232-463-14

Query Match 3.98, Score 36.8, DB 1: Length 7218;
 Best Local Similarity 4.78, Pred No. 1.1e-05

QY 380 TGGAGTCGGGTGGGTGAGCTATGATGACGACCGGTGATTTCTCCATGGGTTTGTGA 439
 Db 1040 TGGCTGACGAGCGAGGAGGAGCTGCGATTTTTTTTTTTTTTTTTTTTTT 1099
 QY 440 GAATCTCTCCACTACTACTGAGCTCTCTCAATCTCAATCTCAATGTTTAAGCTTC 499
 Db 1100 YY 1159
 QY 500 TTCCATTAATCTGCTGACACTGCTCTGAGAGAAAGCTTGGAGCTGATCAGATAA 559
 Db 1160 YY 1219
 QY 560 TGTAGTTCGATCCAGCTGGTGGATTTTGGAAATKACATGATCTCTCTCGATGA 619
 Db 1220 YY 1279
 QY 620 CGGTAGATCATGATCTTCTCAATCAGATCATCATCAGACGACGACGTTCTTTA 679
 Db 1280 YY 1339
 QY 680 TGATCATAGATCCGTAGACAGCTCAGTTCTCTCTAGTACTACTATTAATCTTA 739
 Db 1340 YY 1399
 QY 740 TTTC 743
 Db 1400 YYY 1403

RESULT 3

US-08-332-766A-22
 Sequence 22 Application US/0832766A
 Sequence 2258447
 GENERAL INFORMATION:
 APPLICANT: JEFFREYS, Alec J.
 APPLICANT: ARMOUR, John
 TITLE OF INVENTION: SINGLE TANDEM REPEATS
 NUMBER OF SEQUENCES: 125
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: CUSIMAN DAREY & CUSIMAN, L.L.P.
 STREET: 1100 New York Avenue, N.W.
 CITY: Washington
 STATE: D.C.
 COUNTRY: U.S.A.
 ZIP: 20005-3918
 COMPUTER READABLE FORM:
 HDION TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 APPLICATION DATE: 01-08-1994
 CURRENT APPLICATION NUMBER: US/08/332.766A
 FILING DATE: 01-NOV-1994
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: GB 9326052.9
 FILING DATE: 21-DEC-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: BIRD, Donald J., 25, 393
 REGISTRATION NUMBER: 217211/M94/0434/GB
 TELEPHONE: (202) 861-3000
 TELEFAX: (202) 822-0944
 INFORMATION FOR SEQ ID NO: 22:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 494 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 IMMUNOLOGY SOURCE: near
 CLONE: pT98pt-Fls
 US-08-332-766A-22

Query Match 3.6%; Score 34.2; DB 2; Length 494;
 Best Local Similarity 49.0%; Pred. No. 0.14;
 Matches 72; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

Y 99 GAGATTCGCGACAGACACTGAGACTGTGGTTCGGAACAGAGATCGAGAGACAGTGCAG 158
 Y 156 GAAACGAGAGAAAAG 215
 Y 159 CAAMACAGAGAACCAACGCTGAGAGAGTGGGTGTGCAAGCTGTGAGAGCTGAGAGTCA 218
 Y 216 GAG 275
 Y 219 ATCCGACAG 245
 Y 276 AAGAAATANNNNNNNNNNNNNAATCTCC 302

RESULT 4
 IS-08-485-569-1
 Sequence 1, Application US/08485569
 Patent No. 5679547
 GENERAL INFORMATION:
 APPLICANT: Kivav, Howard C.
 APPLICANT: Samuel, James E.
 APPLICANT: No. 5721115ber, Nils T.
 TITLE OF INVENTION: ADHESIN-OLIGOSACCHARIDE CONJUGATE
 TITLE OF INVENTION: VACCINE FOR HEMOPHILUS INFLUENZAE
 NUMBER OF SOURCES: 2
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Penlite & Edmonds
 STREET: 1155 Avenue of Americas
 CITY: New York
 STATE: New York
 COUNTRY: U.S.A.
 ZIP: 10036
 COMPUTER READABLE FORM:
 MEDIUM TYPE: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08485,569
 FILING DATE: 07-JUN-1993
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/903,079
 FILING DATE: 22-JUN-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/810,966
 FILING DATE: 20-DEC-1991
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/631,698
 FILING DATE: 21-DEC-1990
 ATTORNEY/AGENT INFORMATION:
 NAME: Baldin, Geraldine F.
 REGISTRATION NUMBER: 31,232
 REFERENCE/DOCKET NUMBER: 7969-049
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 669-8864/9741
 TELEFAX: (212) 669-8864/9741
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1511 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: unknown
 MOLECULE TYPE: DNA

IS-08-485-569-1
 Query Match 3.6%; Score 32.8; DB 1; Length 1511;
 Best Local Similarity 50.6%; Pred. No. 0.89;
 Matches 79; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

QY 730 TTAATCTTATTTCAACGAGCGACAAATATACGGGCAATGAGGAATTTGGAGCT 789
 Y 968 TCATCTGATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 1027
 QY 760 ACATCTGAGAAACCCCTAGAAATGAGTACAGGAGTGAGACGAGTACGAGTTTTCGG 849
 DB 1028 AAGTTTTCAGCAATCTCTCTCTGAAAAAGAGAGAGTAAAGCGGGCGATATATACGG 1087
 QY 850 GGAATATCTGTGAGAGATTTCTAGTGTGGCTGCTCAA 885
 DB 1088 CGATGACGCTGCAAAAATCTCAAGTTTCCTGAAA 1123

RESULT 5
 US-08-480-993-1
 Sequence 1, Application US/08480993
 Patent No. 5721115
 GENERAL INFORMATION:
 APPLICANT: Kivav, Howard C.
 APPLICANT: Samuel, James E.
 APPLICANT: No. 5721115ber, Nils T.
 TITLE OF INVENTION: ADHESIN-OLIGOSACCHARIDE CONJUGATE
 TITLE OF INVENTION: VACCINE FOR HEMOPHILUS INFLUENZAE
 NUMBER OF SOURCES: 2
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Penlite & Edmonds
 STREET: 1155 Avenue of Americas
 CITY: New York
 STATE: New York
 COUNTRY: U.S.A.
 ZIP: 10036
 COMPUTER READABLE FORM:
 MEDIUM TYPE: IBM PC compatible
 OPERATING SYSTEM: IBM PC compatible
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08480,993
 FILING DATE: 07-JUN-1993
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/903,079
 FILING DATE: 22-JUN-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/810,966
 FILING DATE: 20-DEC-1991
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/631,698
 FILING DATE: 21-DEC-1990
 ATTORNEY/AGENT INFORMATION:
 NAME: Baldin, Geraldine F.
 REGISTRATION NUMBER: 31,232
 REFERENCE/DOCKET NUMBER: 7969-050
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 790-9090
 TELEFAX: (212) 669-8864/9741
 TELETYPE: 6611 PENITE
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1511 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: unknown
 MOLECULE TYPE: DNA

US-08-480-993-1
 Query Match 3.5%; Score 32.8; DB 1; Length 1511;
 Best Local Similarity 50.6%; Pred. No. 0.89;
 Matches 79; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

QY 730 TTAATCTTATTTCAACGAGCGACAAATATACGGGCAATGAGGAATTTGGAGCT 789

Db 568 TCATGCTGATTAGCAAGGCTTTAATGTAGCGCGCAACAGGTCATTGTGAAGTG 1027
 : Sequence 1, Application US/07903079
 : Patent No. 5843463
 : General Information:
 : APPLICANT: Krivan, Howard C.
 : APPLICANT: Samuel, James E. Nils T.
 : APPLICANT: ADRISIN-OLIGOSACCHARIDE CONJUGATE
 : TITLE OF INVENTION: VACCINE FOR HADMOPIHLUS INFLUENZAE
 : NUMBER OF SEQUENCES: 2
 : COMSEQUENCE ADDRESS: Edmonds
 : ADDRESS: 1155 Avenue of Americas
 : CITY: New York
 : STATE: New York
 : COUNTRY: U.S.A.
 : ZIP: 10036

RESULT 6
 US-07-903-079B-1
 : Sequence 1, Application US/07903079
 : Patent No. 5843463
 : General Information:
 : APPLICANT: Krivan, Howard C.
 : APPLICANT: Samuel, James E. Nils T.
 : APPLICANT: ADRISIN-OLIGOSACCHARIDE CONJUGATE
 : TITLE OF INVENTION: VACCINE FOR HADMOPIHLUS INFLUENZAE
 : NUMBER OF SEQUENCES: 2
 : COMSEQUENCE ADDRESS: Edmonds
 : ADDRESS: 1155 Avenue of Americas
 : CITY: New York
 : STATE: New York
 : COUNTRY: U.S.A.
 : ZIP: 10036

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07903,079B
 FILING DATE: 22-JUN-1992
 PRIORITY DATE: 20-DEC-1991
 APPLICATION NUMBER: US 07/910,966
 FILING DATE: 20-DEC-1991
 APPLICATION NUMBER: US 07/631,698
 FILING DATE: 21-DEC-1990
 ATTORNEY/AGENT INFORMATION:
 NAME: Misrock, S. Leslie
 REFERENCE/DOCKET NUMBER: 7569-004
 TELEPHONE: (212) 790-9090
 TELEFAX: (212) 790-9090
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1611 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: unknown
 MOL-RCLE TYPE: DNA

US-07-903-079B-1
 : Sequence 1, Application US/07903079
 : Patent No. 5843463
 : General Information:
 : APPLICANT: Krivan, Howard C.
 : APPLICANT: Samuel, James E. Nils T.
 : APPLICANT: ADRISIN-OLIGOSACCHARIDE CONJUGATE
 : TITLE OF INVENTION: VACCINE FOR HADMOPIHLUS INFLUENZAE
 : NUMBER OF SEQUENCES: 2
 : COMSEQUENCE ADDRESS: Edmonds
 : ADDRESS: 1155 Avenue of Americas
 : CITY: New York
 : STATE: New York
 : COUNTRY: U.S.A.
 : ZIP: 10036

Oy 850 GGAATATGCTGAAGAGTTTCAGTGTGCTGTACAA 885
 : Sequence 1, Application US/08278091
 : Patent No. 5506139
 : General Information:
 : APPLICANT: LOOSMORE, Sheena M
 : APPLICANT: YAN, Yan-Ping
 : APPLICANT: CHONG, Pele
 : APPLICANT: COHEN, Raymond P.
 : APPLICANT: KATZ, Andrew H.
 : TITLE OF INVENTION: Analog of haemophilus Hln47 Protein with
 : REDUCED PROTEASE ACTIVITY
 : NUMBER OF SEQUENCES: 23
 : COMSEQUENCE ADDRESS: St. Lawrence
 : ADDRESS: Suite 701, 330 University Avenue
 : CITY: Toronto
 : STATE: Ontario
 : COUNTRY: Canada
 : ZIP: M5G 1P9

RESULT 7
 US-08-278-091-1
 : Sequence 1, Application US/08278091
 : Patent No. 5506139
 : General Information:
 : APPLICANT: LOOSMORE, Sheena M
 : APPLICANT: YAN, Yan-Ping
 : APPLICANT: CHONG, Pele
 : APPLICANT: COHEN, Raymond P.
 : APPLICANT: KATZ, Andrew H.
 : TITLE OF INVENTION: Analog of haemophilus Hln47 Protein with
 : REDUCED PROTEASE ACTIVITY
 : NUMBER OF SEQUENCES: 23
 : COMSEQUENCE ADDRESS: St. Lawrence
 : ADDRESS: Suite 701, 330 University Avenue
 : CITY: Toronto
 : STATE: Ontario
 : COUNTRY: Canada
 : ZIP: M5G 1P9

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/278,091
 FILING DATE: 21-JUN-1994
 PRIORITY DATE: 21-JUN-1994
 CLASSIFICATION: 415
 ATTORNEY/AGENT INFORMATION:
 NAME: Stewart, Michael I
 REGISTRATION NUMBER: 24,973
 REFERENCE/DOCKET NUMBER: 7569-038-371
 TELEPHONE: (416) 595-1155
 TELEFAX: (416) 595-1163
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2694 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear

US-08-278-091-1
 : Sequence 1, Application US/08278091
 : Patent No. 5506139
 : General Information:
 : APPLICANT: LOOSMORE, Sheena M
 : APPLICANT: YAN, Yan-Ping
 : APPLICANT: CHONG, Pele
 : APPLICANT: COHEN, Raymond P.
 : APPLICANT: KATZ, Andrew H.
 : TITLE OF INVENTION: Analog of haemophilus Hln47 Protein with
 : REDUCED PROTEASE ACTIVITY
 : NUMBER OF SEQUENCES: 23
 : COMSEQUENCE ADDRESS: St. Lawrence
 : ADDRESS: Suite 701, 330 University Avenue
 : CITY: Toronto
 : STATE: Ontario
 : COUNTRY: Canada
 : ZIP: M5G 1P9

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/278,091
 FILING DATE: 21-JUN-1994
 PRIORITY DATE: 21-JUN-1994
 CLASSIFICATION: 415
 ATTORNEY/AGENT INFORMATION:
 NAME: Stewart, Michael I
 REGISTRATION NUMBER: 24,973
 REFERENCE/DOCKET NUMBER: 7569-038-371
 TELEPHONE: (416) 595-1155
 TELEFAX: (416) 595-1163
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2694 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear

US-08-278-091-1
 : Sequence 1, Application US/08278091
 : Patent No. 5506139
 : General Information:
 : APPLICANT: LOOSMORE, Sheena M
 : APPLICANT: YAN, Yan-Ping
 : APPLICANT: CHONG, Pele
 : APPLICANT: COHEN, Raymond P.
 : APPLICANT: KATZ, Andrew H.
 : TITLE OF INVENTION: Analog of haemophilus Hln47 Protein with
 : REDUCED PROTEASE ACTIVITY
 : NUMBER OF SEQUENCES: 23
 : COMSEQUENCE ADDRESS: St. Lawrence
 : ADDRESS: Suite 701, 330 University Avenue
 : CITY: Toronto
 : STATE: Ontario
 : COUNTRY: Canada
 : ZIP: M5G 1P9

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/278,091
 FILING DATE: 21-JUN-1994
 PRIORITY DATE: 21-JUN-1994
 CLASSIFICATION: 415
 ATTORNEY/AGENT INFORMATION:
 NAME: Stewart, Michael I
 REGISTRATION NUMBER: 24,973
 REFERENCE/DOCKET NUMBER: 7569-038-371
 TELEPHONE: (416) 595-1155
 TELEFAX: (416) 595-1163
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2694 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear

US-08-278-091-1
 : Sequence 1, Application US/08278091
 : Patent No. 5506139
 : General Information:
 : APPLICANT: LOOSMORE, Sheena M
 : APPLICANT: YAN, Yan-Ping
 : APPLICANT: CHONG, Pele
 : APPLICANT: COHEN, Raymond P.
 : APPLICANT: KATZ, Andrew H.
 : TITLE OF INVENTION: Analog of haemophilus Hln47 Protein with
 : REDUCED PROTEASE ACTIVITY
 : NUMBER OF SEQUENCES: 23
 : COMSEQUENCE ADDRESS: St. Lawrence
 : ADDRESS: Suite 701, 330 University Avenue
 : CITY: Toronto
 : STATE: Ontario
 : COUNTRY: Canada
 : ZIP: M5G 1P9

Db 568 TCATGCTGATTAGCAAGGCTTTAATGTAGCGCGCAACAGGTCATTGTGAAGTG 1027
 : Sequence 1, Application US/07903079
 : Patent No. 5843463
 : General Information:
 : APPLICANT: Krivan, Howard C.
 : APPLICANT: Samuel, James E. Nils T.
 : APPLICANT: ADRISIN-OLIGOSACCHARIDE CONJUGATE
 : TITLE OF INVENTION: VACCINE FOR HADMOPIHLUS INFLUENZAE
 : NUMBER OF SEQUENCES: 2
 : COMSEQUENCE ADDRESS: Edmonds
 : ADDRESS: 1155 Avenue of Americas
 : CITY: New York
 : STATE: New York
 : COUNTRY: U.S.A.
 : ZIP: 10036

RESULT 6
 US-07-903-079B-1
 : Sequence 1, Application US/07903079
 : Patent No. 5843463
 : General Information:
 : APPLICANT: Krivan, Howard C.
 : APPLICANT: Samuel, James E. Nils T.
 : APPLICANT: ADRISIN-OLIGOSACCHARIDE CONJUGATE
 : TITLE OF INVENTION: VACCINE FOR HADMOPIHLUS INFLUENZAE
 : NUMBER OF SEQUENCES: 2
 : COMSEQUENCE ADDRESS: Edmonds
 : ADDRESS: 1155 Avenue of Americas
 : CITY: New York
 : STATE: New York
 : COUNTRY: U.S.A.
 : ZIP: 10036

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07903,079B
 FILING DATE: 22-JUN-1992
 PRIORITY DATE: 20-DEC-1991
 APPLICATION NUMBER: US 07/910,966
 FILING DATE: 20-DEC-1991
 APPLICATION NUMBER: US 07/631,698
 FILING DATE: 21-DEC-1990
 ATTORNEY/AGENT INFORMATION:
 NAME: Misrock, S. Leslie
 REFERENCE/DOCKET NUMBER: 7569-004
 TELEPHONE: (212) 790-9090
 TELEFAX: (212) 790-9090
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1611 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: unknown
 MOL-RCLE TYPE: DNA

US-07-903-079B-1
 : Sequence 1, Application US/07903079
 : Patent No. 5843463
 : General Information:
 : APPLICANT: Krivan, Howard C.
 : APPLICANT: Samuel, James E. Nils T.
 : APPLICANT: ADRISIN-OLIGOSACCHARIDE CONJUGATE
 : TITLE OF INVENTION: VACCINE FOR HADMOPIHLUS INFLUENZAE
 : NUMBER OF SEQUENCES: 2
 : COMSEQUENCE ADDRESS: Edmonds
 : ADDRESS: 1155 Avenue of Americas
 : CITY: New York
 : STATE: New York
 : COUNTRY: U.S.A.
 : ZIP: 10036

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07903,079B
 FILING DATE: 22-JUN-1992
 PRIORITY DATE: 20-DEC-1991
 APPLICATION NUMBER: US 07/910,966
 FILING DATE: 20-DEC-1991
 APPLICATION NUMBER: US 07/631,698
 FILING DATE: 21-DEC-1990
 ATTORNEY/AGENT INFORMATION:
 NAME: Misrock, S. Leslie
 REFERENCE/DOCKET NUMBER: 7569-004
 TELEPHONE: (212) 790-9090
 TELEFAX: (212) 790-9090
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1611 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: unknown
 MOL-RCLE TYPE: DNA

US-07-903-079B-1
 : Sequence 1, Application US/07903079
 : Patent No. 5843463
 : General Information:
 : APPLICANT: Krivan, Howard C.
 : APPLICANT: Samuel, James E. Nils T.
 : APPLICANT: ADRISIN-OLIGOSACCHARIDE CONJUGATE
 : TITLE OF INVENTION: VACCINE FOR HADMOPIHLUS INFLUENZAE
 : NUMBER OF SEQUENCES: 2
 : COMSEQUENCE ADDRESS: Edmonds
 : ADDRESS: 1155 Avenue of Americas
 : CITY: New York
 : STATE: New York
 : COUNTRY: U.S.A.
 : ZIP: 10036

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07903,079B
 FILING DATE: 22-JUN-1992
 PRIORITY DATE: 20-DEC-1991
 APPLICATION NUMBER: US 07/910,966
 FILING DATE: 20-DEC-1991
 APPLICATION NUMBER: US 07/631,698
 FILING DATE: 21-DEC-1990
 ATTORNEY/AGENT INFORMATION:
 NAME: Misrock, S. Leslie
 REFERENCE/DOCKET NUMBER: 7569-004
 TELEPHONE: (212) 790-9090
 TELEFAX: (212) 790-9090
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1611 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: unknown
 MOL-RCLE TYPE: DNA

APPLICANT: CHONG, Pele
APPLICANT: CHONG, Raymond P.
APPLICANT: KLEIN, Michel H.
TITLE OF INVENTION: Analog of Haemophilus Hin7 Protein with
Reduced Protease Activity
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:

ADDRESS: Jim & McBurney
STREET: Suite 701, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada

ZIP: MSG IR/
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/483,859
FILING DATE: 07-JUN-1995
CLASSIFICATION: 425

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/296,149
FILING DATE: 26-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/037,001

FILING DATE: 21-JUL-1994
ATTORNEY/AGENT INFORMATION:
NAME: Stewart, Michael I.
REGISTRATION NUMBER: 24,973

TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:
LENGTH: 2894 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

3-483-839-1
Very Match 3.5%; Score 32.8; DB 1; Length 2894;
at Local Similarity 50.6%; Pred. No. 1.3;
Matches 79; Conservative 0; Mismatches 77; Indels 0; Gaps

730 TTATCCTTATTTCACGAGGCAACAATCATACGGACCATTGGAGCAATTTGGAGCT 789
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1655 TCAATCGTGATTTAGCCAAAGCCTTTAATTAAGCGGCACACGCCATTGTGAAGTG 1714

790 ACATGGAGGAAACCTTCAGAAATGGATCAGGAGGTCTCAAGGAGTACCAGGTTTTTTCGG 849
1715 AAGTTTTACCAAAATCTGCTGCTGAAAAAGCAGGACITAAAGCGGCGATATTATCAGG 1774

1775 CGATGACGGTCAAAAAATCTCAAGTTTCGCTGAA 1810

8-472-173-1
Sequence 1, Application US/08472173
Rent No. 565353
GENERAL INFORMATION:

APPLICANT: LOOSMORE, Sheila M
APPLICANT: YANG, Yan-Ping
APPLICANT: CHONG, Pele
APPLICANT: OOMEN, Raymond P.
APPLICANT: KLEIN, Michel H.

TITLE OF INVENTION: Analog of Haemophilus Hin4 / Protein with
TITLE OF INVENTION: Reduced Protease Activity
NUMBER OF SEQUENCES: 23

REFERENCE/DOCKET NUMBER: 1038-390
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2894 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-801-499-1

Query Match 3.5%: Score 32.8; DB 2; Length 2894;

Best Local Similarity 50.8%: Pred. No. 1.3; Mismatches 77; Indels 0; Gaps 0;
Matches 79; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

y 730 TTAATCTTTATTCAACGAGGACCAAAATCATACGGGACCAATGAGGAAATTTGGAGCT 789
b 1655 TCAATGCTGATTATAGCGAAGAGCTTTATGTAGCGCGGACACAGGCGCATTTGTAGTG 1714
y 790 ACATGAGGAAAGCACTTGAAGATGATCAGGAGGTGTGAAGGATGACGATTTTTCGGS 849
b 1715 AGATTTCAGCAAAATCTGCTGTGTAAGGAGCGAGCTTAAAGCGGCGCATATTATCAGG 1774
y 850 GGAATATGCTGAAGAGATTTCAGTGTGCGTACAA 885
b 1775 CGATGACGCTCAAAAATCTCAAGTTTCGCTGAAA 1810

RESULT 13

S-08-801-499-1

Sequence 1. Application US/08801499

Patent No. 5962430

Genetic Information:

APPLICANT: LOOSMORE, Shreeta M.

APPLICANT: YANG, Yan-Ping

APPLICANT: CHONG, Pele

APPLICANT: COMEN, Raymond P.

APPLICANT: KLEIN, Michael B.

TITLE OF INVENTION: ANALOG OF HAEMOPHILUS HIN47 Protein with Reduced Protease Act

NUMBER OF SEQUENCES: 23

CORRESPONDENCE ADDRESS:

ADDRESSEE: Sim & McBurney

STREET: 6th Floor, 330 University Avenue

CITY: Toronto

STATE: Ontario

COUNTRY: Canada

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

CURRENT APPLICATION DATA: Patent Release #1.0, Version #1.25

APPLICATION NUMBER: US/08/801,499

FILING DATE:

CLASSIFICATION: 514

PRIOR APPLICATION NUMBER:

APPLICATION NUMBER: US 08/482,816

FILING DATE: 07-JUN-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/296,149

FILING DATE: 07-JUN-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/278,091

FILING DATE: 21-JUL-1994

ATTORNEY/AGENT INFORMATION:

REGISTRATION NUMBER: 24,973

REFERENCE/DOCKET NUMBER: 1038-671 MIS:jb

TELECOMMUNICATION INFORMATION:

TELEPHONE: (416) 595-1135

TELEFAX: (416) 595-1163

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2894 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-801-499-1

Query Match 3.5%: Score 32.8; DB 2; Length 2894;

Best Local Similarity 50.8%: Pred. No. 1.3; Mismatches 79; Indels 0; Gaps 0;
Matches 79; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

y 730 TTAATCTTTATTCAACGAGGACCAAAATCATACGGGACCAATGAGGAAATTTGGAGCT 789
b 1655 TCAATGCTGATTATAGCGAAGAGCTTTATGTAGCGCGGACACAGGCGCATTTGTAGTG 1714
y 790 ACATGAGGAAAGCACTTGAAGATGATCAGGAGGTGTGAAGGATGACGATTTTTCGGS 849
b 1715 AGATTTCAGCAAAATCTGCTGTGTAAGGAGCGAGCTTAAAGCGGCGCATATTATCAGG 1774
y 850 GGAATATGCTGAAGAGATTTCAGTGTGCGTACAA 885
b 1775 CGATGACGCTCAAAAATCTCAAGTTTCGCTGAAA 1810

RESULT 14

US-08-615-271-1

Sequence 1. Application US/08615271

Patent No. 5918101

Genetic Information:

APPLICANT: LOOSMORE, Shreeta M.

APPLICANT: YANG, Yan-Ping

APPLICANT: CHONG, Pele

APPLICANT: COMEN, Raymond P.

APPLICANT: KLEIN, Michael B.

TITLE OF INVENTION: ANALOG OF HAEMOPHILUS HIN47 WITH REDUCED

TITLE OF INVENTION: PROTEASE ACTIVITY

NUMBER OF SEQUENCES: 23

CORRESPONDENCE ADDRESS:

ADDRESSEE: Sim & McBurney

STREET: 6th Floor, 330 University Avenue

CITY: Toronto

STATE: Ontario

COUNTRY: Canada

ZIP: M5G 1R7

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

CURRENT APPLICATION DATA: Patent Release #1.0, Version #1.30

APPLICATION NUMBER: US/08/615,271

FILING DATE: 20-JUN-1996

CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:

NAME: Stewart, Michael I.

REGISTRATION NUMBER: 24,973

REFERENCE/DOCKET NUMBER: 1038-560

TELECOMMUNICATION INFORMATION:

TELEPHONE: (416) 595-1135

TELEFAX: (416) 595-1163

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 2894 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MEDIUM TYPE: DNA (genomic)

US-08-615-271-1

Query Match 3.5%: Score 32.8; DB 2; Length 2894;

Best Local Similarity 50.8%: Pred. No. 1.3; Mismatches 79; Indels 0; Gaps 0;
Matches 79; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

4 nucleic - nucleic search, using sw model

on on: January 25, 2003, 22:30:57 ; Search time 63 Seconds

file: US-09-701-023_1_COPY_01_1024

sequence: 1 tggcgactctctctctc.....atggctcgtgaagctttaa 944

Sorting table: IDENTITY_MUC

Gapop 10.0%, Gapext 1.0

searched: 396772 seqs. 224632407 residues

total number of hits satisfying chosen parameters: 793544

Minimum DB seq length: 0

Minimum DB seq length: 2000000000

FAST-processing: Minimum Match 0%

Minimum Match 100%

Listing first 45 summaries

- Database : PublishedApplications.NA.*
- 1: /cgn2.6/ptodata/pub/pubseq.seq.*
 - 2: /cgn2.6/ptodata/pub/pubseq.seq.*
 - 3: /cgn2.6/ptodata/pub/pubseq.seq.*
 - 4: /cgn2.6/ptodata/pub/pubseq.seq.*
 - 5: /cgn2.6/ptodata/pub/pubseq.seq.*
 - 6: /cgn2.6/ptodata/pub/pubseq.seq.*
 - 7: /cgn2.6/ptodata/pub/pubseq.seq.*
 - 8: /cgn2.6/ptodata/pub/pubseq.seq.*
 - 9: /cgn2.6/ptodata/pub/pubseq.seq.*
 - 10: /cgn2.6/ptodata/pub/pubseq.seq.*
 - 11: /cgn2.6/ptodata/pub/pubseq.seq.*
 - 12: /cgn2.6/ptodata/pub/pubseq.seq.*
 - 13: /cgn2.6/ptodata/pub/pubseq.seq.*
 - 14: /cgn2.6/ptodata/pub/pubseq.seq.*

prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	35.8	3.8	465237	10	US-09-933-267A-1/c
2	34.6	3.7	22669	9	US-09-938-920-113
3	34.2	3.6	389	9	US-09-756-692-2664
4	34	3.6	470	10	US-09-864-761-701
5	33.6	3.5	325	10	US-09-867-701-7516
6	33.5	3.5	325	10	US-09-867-701-7516
7	33	3.5	331	10	US-09-864-761-2932
8	33	3.5	365	10	US-09-764-864-338
9	33	3.5	511	10	US-09-864-761-12665
10	32.8	3.4	389	9	US-09-867-701-7516
11	32.8	3.4	389	9	US-09-867-701-7516
12	32.4	3.4	389	10	US-09-864-761-10334
13	32.4	3.4	470	10	US-09-864-761-1484
14	32.4	3.4	470	10	US-09-864-761-15622
15	32.2	3.4	1103	10	US-09-864-761-15622
16	32.2	3.4	1103	10	US-09-864-761-15622
17	32.2	3.4	1107	9	US-09-864-921-84
18	32.2	3.4	1155	10	US-09-815-242-8511
19	32.2	3.4	3111	10	US-09-728-721-56

- c 20 32.2 3.4 4244 10 US-09-728-721-54
- c 21 32 3.4 2000 9 US-09-938-8424-2854
- c 22 32 3.4 37768 10 US-09-070-927A-100
- c 23 31.6 3.3 148567 10 US-09-801-8768-3
- c 24 31.6 3.3 148567 10 US-09-801-8768-3
- c 25 31.4 3.3 270 10 US-09-294-0938-4115
- c 26 31.4 3.3 300 10 US-09-294-0938-4388
- c 27 31.4 3.3 13605 10 US-09-764-847-1461
- c 28 31.4 3.3 13605 10 US-09-764-847-1461
- c 29 31.4 3.3 32768 10 US-09-070-927A-16
- c 30 31.2 3.3 3356 9 US-09-756-692-8787
- c 31 31.2 3.3 360004 10 US-09-949-654-3
- c 32 31.2 3.3 7805 12 US-10-044-090-370
- c 33 31 3.3 7805 12 US-10-044-090-370
- c 34 31 3.3 11872 10 US-09-764-847-1390
- c 35 31 3.3 152331 9 US-10-093-407-16
- c 36 30.8 3.3 2851 10 US-09-764-855-301
- c 37 30.8 3.3 2851 10 US-09-764-855-302
- c 38 30.8 3.3 2851 10 US-09-764-855-302
- c 39 30.8 3.3 174424 10 US-09-967-768A-314
- c 40 30.6 3.2 174453 10 US-09-804-471A-3
- c 41 30.6 3.2 584 10 US-09-807-750-5214
- c 42 30.4 3.2 584 10 US-09-807-750-5214
- c 43 30.4 3.2 1323 9 US-09-910-186A-3
- c 44 30.4 3.2 1326 9 US-09-910-186A-5
- c 45 30.4 3.2 1332 9 US-09-910-186A-1

ALIGNMENTS

RESULT 1

US-09-933-267A-1/c

Sequence 1, Application US/09933267A

Patent No. US2002123095A1

GENERIC INFORMATION: Francis et al

TITLE OF INVENTION: Estrogen receptor alpha variants and

FILE REFERENCE: C1000258C14 US/09/933-267A

CURRENT FILING DATE: 2001-08-21

PRIOR FILING DATE: 2001-08-21

PRIOR FILING DATE: 1999-10-20

PRIOR APPLICATION NUMBER: 60/183756

PRIOR APPLICATION NUMBER: 08/592414

PRIOR FILING DATE: 2000-10-20

PRIOR APPLICATION NUMBER: 09/768184

PRIOR FILING DATE: 2001-01-24

PRIOR APPLICATION NUMBER: 08/604076

PRIOR FILING DATE: 2001-03-13

PRIOR APPLICATION NUMBER: 09/826314

PRIOR FILING DATE: 2001-04-05

NUMBER OF SEQ ID NOS: 2

SEQ ID NO 1: 1

SEQ ID NO 1: 1

SEQ ID NO 1: 1

SEQ ID NO 1: 1

SEQ ID NO 1: 1

SEQ ID NO 1: 1

SEQ ID NO 1: 1

SEQ ID NO 1: 1

SEQ ID NO 1: 1

SEQ ID NO 1: 1

SEQ ID NO 1: 1

[illegible]

[illegible][illegible]

This Page Blank (uspto)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

M protein - protein search, using sw model

Run on: January 25, 2003, 22:37:22 ; Search time 70 Seconds

File: US-09-701-023-4
Sequence: 1 MATSLPMSFSDNSGVNDP.....TTSSLVGDCSPNTDLSKL 314
597,724 Million cell updates/sec

Scoring table: BLOSUM62

Gap: 10.0, Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database: A_Genesc101002.*

```
1: /SID82/gcgdata/genescq/genescsp-emb1/AA1980.DAT.*
2: /SID82/gcgdata/genescq/genescsp-emb1/AA1981.DAT.*
3: /SID82/gcgdata/genescq/genescsp-emb1/AA1982.DAT.*
4: /SID82/gcgdata/genescq/genescsp-emb1/AA1983.DAT.*
5: /SID82/gcgdata/genescq/genescsp-emb1/AA1984.DAT.*
6: /SID82/gcgdata/genescq/genescsp-emb1/AA1985.DAT.*
7: /SID82/gcgdata/genescq/genescsp-emb1/AA1986.DAT.*
8: /SID82/gcgdata/genescq/genescsp-emb1/AA1987.DAT.*
9: /SID82/gcgdata/genescq/genescsp-emb1/AA1988.DAT.*
10: /SID82/gcgdata/genescq/genescsp-emb1/AA1989.DAT.*
11: /SID82/gcgdata/genescq/genescsp-emb1/AA1990.DAT.*
12: /SID82/gcgdata/genescq/genescsp-emb1/AA1991.DAT.*
13: /SID82/gcgdata/genescq/genescsp-emb1/AA1992.DAT.*
14: /SID82/gcgdata/genescq/genescsp-emb1/AA1993.DAT.*
15: /SID82/gcgdata/genescq/genescsp-emb1/AA1994.DAT.*
16: /SID82/gcgdata/genescq/genescsp-emb1/AA1995.DAT.*
17: /SID82/gcgdata/genescq/genescsp-emb1/AA1996.DAT.*
18: /SID82/gcgdata/genescq/genescsp-emb1/AA1997.DAT.*
19: /SID82/gcgdata/genescq/genescsp-emb1/AA1998.DAT.*
20: /SID82/gcgdata/genescq/genescsp-emb1/AA1999.DAT.*
21: /SID82/gcgdata/genescq/genescsp-emb1/AA2000.DAT.*
22: /SID82/gcgdata/genescq/genescsp-emb1/AA2001.DAT.*
23: /SID82/gcgdata/genescq/genescsp-emb1/AA2002.DAT.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1628	99.6	314	21	Arabidopsis thaliana
2	1628	99.6	314	21	Arabidopsis thaliana
3	103	6.3	1593	22	Arabidopsis thaliana
4	102	6.2	430	22	Arabidopsis thaliana
5	102	6.2	430	22	Arabidopsis thaliana
6	102	6.2	430	22	Arabidopsis thaliana
7	102	6.2	430	22	Arabidopsis thaliana
8	101.5	6.2	729	22	Arabidopsis thaliana
9	100	6.1	1059	22	Arabidopsis thaliana
10	99.5	6.0	607	22	Arabidopsis thaliana

Drosophila melanogaster
Arabidopsis thaliana
Drosophila melanogaster
Drosophila melanogaster
Novel human diageno
Drosophila melanogaster
Drosophila melanogaster
Drosophila melanogaster
S. cerevisiae appt
S. cerevisiae pacc
Arabidopsis thaliana
Arabidopsis thaliana
Arabidopsis thaliana
Human SRC-3 protei
Maize cinnamyl alc
Mouse ischaemic co
Human EMS protein
Novel human diageno
Drosophila melanogaster
Drosophila melanogaster
Drosophila melanogaster
Candida albicans C
Novel human diageno
Human transcriptio
Human transcriptio
Human transcriptio
Drosophila melanogaster
Drosophila melanogaster
Soybean P418 iron
ER interacting dom
ABI (Amplified in

ALIGNMENTS

RESULT 1
ID: AAB26111 standard; Protein: 314 AA.
AC AAB26111:
AC AC
DE 30-JAN-2001 (first entry)
DE Arabidopsis thaliana sporocytless spl protein.
XX Sporocytless; spl: malocyte formation; plant sterility; seedless fruit;
XX pollenless flower; ds element.
XX Arabidopsis thaliana.
XX Arabidopsis thaliana.

Key Location/Qualifiers
FT Misc-difference 23
FT /label= unknown
FT /note= *shown as Lrg. encoded by AGR*

W0200055907-AL

28-SEP-2000.

22-MAR-1999; 99NO-SG00023.

22-MAR-1999; 99NO-SG00023.

(MOLE-) INST MOLECULAR AGROBIOLOGY.

Ye D, Yang W, Sundaresan V, Xu J;

WPI; 2000-594578/56.

X A (PEKE) PE CORP NY.
X Z
X Z Venter JC, Adams M, Li PHD, Myers BW.
X R WPI; 2001-656860/75.
X R N-PSDB; ABL0385.
X T New isolated nucleic acid detection reagent for detecting 1000 or more
X T genes from Drosophila and for elucidating cell signalling and cell-cell
X T interactions -
X Z
X Z Disclosure; SEQ ID NO 19538; 21pp + Sequence Listing; English.
X C The invention relates to an isolated nucleic acid detection reagent
X C capable of detecting 1000 or more genes from Drosophila. The invention is
X C useful in developmental biology and in elucidating cell signalling and
X C cellular interactions. The invention also relates to a method for detecting
X C insecticides, therapeutics and pharmaceutical drugs. The invention
X C discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
X C sequences (ABL01840-ABL616175) and the encoded proteins
X C (ABL16176-ABL30511).
X C The sequence data for this patent did not form part of the printed
X C specification, but was obtained in electronic format directly from WIPO
X C at ftp://ipo.int/pub/published_pct_sequences.

X Z Sequence 1693 AA:
Query Watch
Best Local Similarity 6.34; Score 103; DB 22; Length 1693;
Matches 77; Conservative 44; Mismatches 131; Indels 108; Gaps 15;
7 FMTDQKNSGNOLLRLNRLVNSSTSEITRLKSGRK-----PGSKTGQGXQ 56
b 964 FLSTVSTYS-TSTSTVTSASASASASASNSGRLOQTVPILSPSNMERRK 1022
y 57 KPTLENGVAKLEPRTTEKKQLAAANTVDTSVSASINATLPPVDPVULQGP 116
b 1023 LKSHLKLLQGGGQGEAGRTGTGAGAVMTGSKVNSASCKGQ-GLPAEP--TLGPP 1079
y 117 SS-----LGNRYG-QVSGQVMDVPISFWQVETSTHLSLSNQ 162
b 1080 ASAKEESFPFGAAAGSVNNAQNGNGSGVS-----GNGSGTVNDSSA---- 1126
y 163 MFASNNNRDQCFKKKLLQGGQNNVNSNGQFSKYTMIPPMNGDQDTLLASDHORS 222
b 1127 -NHSSE-----PSNN-----GNSATCTPPT-----S 1147
y 223 QGTLVDIERIAAAS-----VSASSTTIPFNKAVNHTGMEF-----GSYVE 266
b 1148 WAPVPPPIAPASGVPLPFGVNSVPTLITVPTFNSQSFAPNAGLGGHGT 1207
y 267 GPNRMSGVKYEPPFKYGRVRSVWATTS--SLWGDGS-----PRTDLKAL 314
b 1208 GRLVGVGVSANLPGAGASASASASGTSVNSVLSASKNLHSHDLMAL 1267

RESULT 4
BG13205
X A ASG13205 standard; Protein; 430 AA.
X C ASG13205;
X X
X X 18-FEB-2002 (first entry)
X T Novel human diagnostic protein #13196.
X M human; chromosome mapping; gene mapping; gene therapy; forensic;
X M food supplement; medical imaging; diagnostic; genetic disorder.
X S Homo sapiens.
X N W0200175067-A2.

XX PD 11-OCT-2001.
XX XZ
XX XZ 30-MAR-2001; 2001WO-US08631.
XX PR 31-MAR-2000; 2000US-0540217.
XX PR 23-AUG-2000; 2000US-0649167.
XX XZ (HY3E-) HYSCO INC.
XX PI Dmanac RT, Liu C, Tang YT;
XX XZ WPI; 2001-639162/73.
XX R N-PSDB; ASG77352.
XX PD New isolated polynucleotide and encoded polypeptides, useful in
XX PD diagnostics, forensics, gene mapping, identification of mutations
XX PD and for genetic disorders or other traits and to assess
XX PD biodiversity.
XX Claim 20; SEQ ID NO 43564; 103pp; English.
XX CC The invention relates to isolated polynucleotide (i) and
XX CC polypeptide (ii) sequences. (i) is useful as hybridisation probes,
XX CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX CC gene mapping, and in recombinant production of (ii). The
XX CC polynucleotide sequences and in diagnostics and expressed sequence tags
XX CC for identifying altered genes. (ii) is useful in various techniques
XX CC to restore normal activity of (ii) or to treat disease states involving
XX CC quantitating a polypeptide in tissue, as molecular weight markers and as
XX CC imaging of sites expressing (ii) and (ii) are useful for treating
XX CC disorders involving aberrant protein expression or biological activity.
XX CC The polypeptide and polynucleotide sequences have applications in
XX CC diagnostics, forensics, gene mapping, identification of mutations
XX CC and for genetic disorders or other traits and to produce data on DNA and
XX CC amino acid sequences. Abc00010-ABG30377 represent novel human
XX CC diagnostic amino acid sequences of the invention.
XX CC Note: the sequence data for this patent did not appear in the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp://ipo.int/pub/published_pct_sequences.

XX Z Sequence 430 AA:
Query Watch
Best Local Similarity 6.24; Score 102; DB 22; Length 430;
Matches 61; Conservative 39; Mismatches 94; Gaps 14;
y 35 KVTETLKSERKSGHSTQCKKKPTLEKGVAKLEPRTTEKKQLAAANTVDTSVAS 94
Db 70 TRUKIMERNRRGT-----RTEKARNEVDRTVFTNIGS--AVAA 108
y 95 TSNRRPLPDPVGVVLQPFSGESGNRYCGVSGQVMDVPISFWQVETSTHIE 154
Db 109 AAADQVTSNPVTP-----HTFRHSTAMHLVAG-----IPLKVLQSLMGHKSISPT-- 155
y 155 LSTSLINSPHFNNSNRD--TCF-----KKRLDGDNNVNSNGGFS 197
Db 156 EYVTKVALLVARRHCKNLKCTHSUGVMAETAVINHKKNSSP--IVQSDNLTEA 212
y 198 KYTIMPPMNGYDQ-----YLQSDHQRHSGQFLYDH-----RIARAASVSSASTTPE 247
Db 213 NYSL-----SDKKRMULTF-VQIRSDQTLQDHGICELIVAKTAIFG----- 257
y 248 FNEATNHTGMEEGSGMEEN--PENGSGGVKYEFPF 283
Db 258 ---LTSASASTEPREGSNTFRPEDAGDEKTESPP 292
RESULT 5
ABG13382

OY 6 FPMATDSNGVN-PNDLNTLRILVNSGCEIFRTETLKSGCRPGSKTQOKKKPTLRQ 64
 DB 112 PLVSLNGLWVSLKPEPT-VOLICNSFGSIRMQIAENY---YDSIDSDENAHATLRQ 165
 OY 65 OY-----AKLEHOREEKKQIAAATNODTSVATSNATRLPVPDGVVLQGV 110
 DB 166 ILAASLRHKPRFPAHKAHLRLNSPKKLMPPEAEPHKSSS-PDPTLKPVLDP?-- 222
 OY 111 VQGFPSVSSASTTINPKYCGVSGCWMDPTISPFGVFETSTTHLSLSTSNPQFNASSN- 169
 DB 223 -----PLTANLSKAADVANVSANVANGS--AALECRWIVGLGTETMPKQISUSKSL 273
 OY 170 NRQDTCFKKRRLDGDNNVVRNSGSGFSKTYMTPPMPNCTDQYLOSDDHQRSGCFLYD 229
 DB 274 QTVASCTKLRGKRPPTV-----LACTP-----EELSDAGS-----SGGLATY 315
 OY 230 RIARAASVSSASTTINPKYCGVSGCWMDPTISPFGVFETSTTHLSLSTSNPQFNASSN- 169
 DB 316 SKRMFNVSSSESENFPLVTSVAVNQILER-----ETEPSAGS 356
 RESULT 9
 ABB71639
 AA ABB71639 standard; Protein: 1059 AA.
 AC ABB71639;
 XX XX
 XX 26-MAR-2002 (first entry)
 XX XX
 XX 26-MAR-2002 (first entry)
 XX XX
 XX Drosophila melanogaster polypeptide SEQ ID NO 41709.
 XX Drosophila; developmental biology; cell signalling; insecticide;
 XX pharmaceutical.
 XX Drosophila melanogaster.
 XX Drosophila melanogaster.
 XX W0200171042-A2.
 XX 27-SEP-2001.
 XX 23-MAR-2001; 2001WO-US09231.
 XX 23-MAR-2001; 2001WO-US09231.
 XX 21-MAR-2000; 2000US-191637P.
 XX 11-JUL-2000; 2000US-0614150.
 XX (PEKE) PE CORP NY.
 XX Venter JC, Adams M, Li PWD, Myers EN.
 XX WPI; 2001-656860/75.
 XX N-PSDB; ABL15742.
 XX New isolated nucleic acid detection reagent for detecting 1000 or more
 XX genes from Drosophila and for elucidating cell signalling and cell-cell
 XX interactions.
 XX Disclosures: SEQ ID NO 41709; Zipp + Sequence Listing; English.
 XX The invention relates to an isolated nucleic acid detection reagent
 XX capable of detecting 1000 or more genes from Drosophila. The invention is
 XX useful in developmental biology and in elucidating cell signalling and
 XX cell-cell interactions in higher eukaryotes for the development of
 XX insecticides, therapeutics and pharmaceutical drugs. The invention
 XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 XX sequences (ABL01840-ABL16175) and the encoded proteins
 XX (AB057373-AB072072).
 XX The invention relates to this patent did not form part of the printed
 XX specification, but was obtained in electronic format directly from WIPO
 XX at ftp.wipo.int/pub/published_pct_sequences.
 XX Sequence 1059 AA;
 XX 6.1%; Score 100; DB 22; Length 1059;
 XX Query Match:

Best Local Similarity 17.5%; Pred. No. 1.5;
 Matches 95; Conservative 45; Mismatches 120; Indels 70; Gaps 9;
 OY 8 MSTQDSNGVNNDLIP-NTFLVNSGCEIFRTETLKSGCRPGSKTQOKKKPTLRQ 66
 DB 605 MEQOQHPSPNSNRRGSSSLAGTSGNKKDKRSTVSNSSSSGYETQO----- 655
 OY 67 AKLEHOREEKKQIAAATNODTSVATSNATRLPVPDGVVLQGVLPSSGASGNEIYC 126
 DB 655 -----STTNGSNSSASSTT-----PFLAASFYPAFTSGS----- 692
 OY 127 GVGSGVMDLIVISPMGCVETSTTHLSLSTSNPQFNASSNRCETCFKKRL----- 181
 DB 693 -ASGTSKRPMPQDSSTVSYKSNNAQVYASN-----SKRMGSGVP 732
 OY 182 -DGDNNVVRNSGSGFSKTYMTPPMPNCTDQYLOSDDHQRSGCFLYDIRAASVNAS 240
 DB 733 SGSGNQRQSRASGNSRGTQVPPPSNRSRSTQQNHQKQ-----VQCKQAFSQ 785
 OY 241 STTINPKYCGVSGCWMDPTISPFGVFETSTTHLSLSTSNPQFNASSNRCETCFKKRL----- 181
 DB 786 QOQOQOQTHQOKHHPSPSQOAAAAAHAYAHATDTSATPTDF 830
 RESULT 10
 ABB61226
 AA ABB61226 standard; Protein: 607 AA.
 AC ABB61226;
 XX XX
 XX 26-MAR-2002 (first entry)
 XX XX
 XX Drosophila melanogaster polypeptide SEQ ID NO 10470.
 XX Drosophila; developmental biology; cell signalling; insecticide;
 XX pharmaceutical.
 XX Drosophila melanogaster.
 XX W0200171042-A2.
 XX 27-SEP-2001.
 XX 23-MAR-2001; 2001WO-US09231.
 XX 23-MAR-2000; 2000US-191637P.
 XX 11-JUL-2000; 2000US-0614150.
 XX (PEKE) PE CORP NY.
 XX Venter JC, Adams M, Li PWD, Myers EN.
 XX WPI; 2001-656860/75.
 XX N-PSDB; ABL05329.
 XX New isolated nucleic acid detection reagent for detecting 1000 or more
 XX genes from Drosophila and for elucidating cell signalling and cell-cell
 XX interactions.
 XX Disclosures: SEQ ID NO 10470; Zipp + Sequence Listing; English.
 XX The invention relates to an isolated nucleic acid detection reagent
 XX capable of detecting 1000 or more genes from Drosophila. The invention is
 XX useful in developmental biology and in elucidating cell signalling and
 XX cell-cell interactions in higher eukaryotes for the development of
 XX insecticides, therapeutics and pharmaceutical drugs. The invention
 XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 XX sequences (ABL01840-ABL16175) and the encoded proteins
 XX (AB057373-AB072072).
 XX The invention relates to this patent did not form part of the printed
 XX specification, but was obtained in electronic format directly from WIPO
 XX at ftp.wipo.int/pub/published_pct_sequences.
 XX Sequence 1059 AA;
 XX 6.1%; Score 100; DB 22; Length 1059;
 XX Query Match:

Drosophila melanogaster.

WO200171042-A2.

27-SEP-2001.

23-MAR-2001: 2001WO-US09231.

23-MAR-2000: 2000US-191637P.

11-JUL-2000: 2000US-0614150.

(PEKE) PE CORP NV.

Venter JC, Adams M, Li PWD, Myers BW:
WPI: 2001-656860/75.
N-PSDB: ABL05853.

New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signaling and cell-cell interactions

Disclosure: SEQ ID NO 12060; 21pp + Sequence Listing; English.

The invention relates to an isolated nucleic acid detection reagent that can be used in developmental biology and in elucidating cell signaling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genes and their sequences (ABL01676-ABL0511), expressed DNA (AB057737-AB072072), and the encoded proteins (AB057737-AB072072).

The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at: http://wipo.int/pub/published_pat_sequences.

Sequence 704 AA:

Query Match	5.94	Score 96.5	DB 22	Length 704
Seq. Local Similarity	22.14	Pred. No. 18		
Matches	53	Conservative 33	Mismatches 79	Indels 75
				Gaps 11
87	GPSVSIVASISNNKTRLPVPVGGVGLQFPSS	---	LGSRNYTCGVSGVSGQWDPVI	---
88	GPSVSIVASISNNKTRLPVPVGGVGLQFPSS	---	LGSRNYTCGVSGVSGQWDPVI	---
25	GSISASASAVSYSSISCTAV	---	GGTILSGAKSPCLASSTSVIVENAGSMDLIDQSL	82
141	---	SPWQFV	---	ETESTETHEL
83	HKHSKISVSGASGAGGASFTSLNLLQWKNK	GLVGVKDSISLSSSPQPMHAGAS	---	138
174	TCFKKKRLDQ	---	QNNYVNSG	---
139	---	LFSGDENVSLRNFMGAGFNRKSCGG	PLPLKMLFKQLHQCGHQHQHQHQ	191
221	RGQFTYVHHAIRADASVASSTFTIPFNATN	QWPEW	---	275
152	HQKQDQQLHQBQ	---	QQLSPSLALHHHQHQDQL	240

RESULT 15
 AG13245
 D AG13249 standard; Protein: 1098 aa.
 K
 C AG13249;
 C
 F 18-FEB-2002 (first entry)
 K
 E Novel human diagnostic protein #13240.
 K
 N Human; chromosome mapping; gene mapping; gene therapy; formaic;
 N food supplement; medical imaging; diagnostic; genetic disorder.
 N Homo sapiens.
 N

XX	PM	WO200175067-A2.
XX	PP	11-OCT-2001.
XX	PP	30-MAR-2001; 2001NO-US008631.
XX	PP	31-MAR-2000; 2000US-0540217.
XX	PP	23-AUG-2000; 2000US-0649167.
XX	PP	(HYSE-) HYSEQ INC.
XX	PP	Dmanac RT, Liu C, Tang YT;
XX	PP	NPT, 2001-630562/73.
XX	PP	USPTO, 2001-0630562/73.
XX	PP	NCSD, AAS77416.
XX	PP	New isolated polynucleotide and encoded polypeptides, useful in
XX	PP	diagnostics, forensics, gene mapping. Identification of mutations
XX	PP	useful for genetic disorders or other traits and to assess
XX	PP	biodiversity
XX	PP	Claim 20; SEQ ID NO 43608; 103pp; English.
XX	PP	The invention relates to isolated polynucleotide (I) and
XX	PP	polypeptide (II) sequences. (I) is useful as hybridisation probes.
XX	PP	polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX	PP	gene mapping, and in recombinant production of (II). The
XX	PP	polypeptides are also used in diagnostics as expressed sequence tags
XX	PP	to identify genes and to identify mutations. The invention also
XX	PP	restores normal activity of (II) or to treat disease states involving
XX	PP	(II). (I) is useful for generating antibodies against it, detecting or
XX	PP	quantitating a polypeptide in tissue, as molecular weight markers and as
XX	PP	immunological reagents for binding assays. (II) is useful for treating
XX	PP	leaving of sites expressing (II) (I) and (II) are useful for treating
XX	PP	disorders involving aberrant protein expression or biological activity.
XX	PP	The polypeptide and polynucleotide sequences have applications in
XX	PP	diagnostics, forensics, gene mapping. Identification of mutations
XX	PP	useful for genetic disorders or other traits and to assess biodiversity
XX	PP	and to produce other types of data and products dependent on DNA and
XX	PP	amino acid sequences. A860010-A8630377 represent novel human
XX	PP	diagnostic amino acid sequences of the invention.
XX	PP	Note: The sequence data for this patent did not appear in the printed
XX	PP	version of the patent document. The sequence data format differed from WIPO
XX	PP	at: ftp://wipointel.wipo.int/pub/published/pct_sequences

Query Match 5.9%; Score 96.5; DB 22; Length 1098;
Best Local Similarity 21.3%; Pred. No. 3.5;
Matches 66; Conservative 47; Mismatches 98; Indels 99; Gaps 17;

OY	11	DNKSNQ-----NOLLEWTH-----DYVNSGIFRETTLSKHSRQSCQYOKK 58
DB	59	DECGWGGPYTPVNNALNNKDKITDYLGLSPVFRKDLMBNRNRTG-----819
OY	59	PTLRNGWAKLORIEEKKQAAATVGTSSVSSISNNAETLPPVDFGVVLQGFSS 118
OY	77	RTKMLINVTBYVNTVTS-----RANAGAGGVSVFVPP-----HTFMS 664
OY	820	-----RTEKMLINVTBYVNTVTS-----RANAGAGGVSVFVPP-----HTFMS 664
OY	113	LASNRIVYCGSGOWIDVPSIPSGFTSTETLSLSINPQNSMNRCD-----TC 175
DB	865	YAMHUYLAG-----TFLVYQSLKMSHSST-----EYTVKFLVDAWNRCDMLTC 914
OY	176	F-----KKKLELDGQNVNNGGSGSKYTIMPPPPNKD-----XLYLSD 217
DB	913	HTSHSGVWARTVNIKKKNSPR-----IVGNSLLEAYSL-----SEQKRELYLF-VD 965
OY	218	HQRSGFLYD-----RIARAASV-----SASSTT-----INPYEATNTHTGPEEFSY 264
OY	266	QIKNSQDQTHQDHCIGELIWTATVIGLISIRANVKEITNY-----ANLUSY 1016
OY	265	MECPNPGSG 774

DB 1017 C0YKPDSC 1026

Search completed: January 26, 2003, 00:03:36
Job time : 78 secs

: LENGTH: 656 amino acids
 : Domain(s):
 : TOPOLAGY: 1 linear
 : MOLECULE TYPE: protein
 us-08-343-443B-2

Query Match 5.78; Score 93; DB 2; Length 656;
 Best Local Similarity 20.28; Pred. No. 0.5; 127; Indels 104; Gaps 13;

Matches 67; Conservative 34; Mismatches 104; Gaps 13;

Oy 16 GPNDLRLRVVWSSGEITFTTLKGRKP---GSKTQKQKKPPLRGKGVAKLEQ 72

Db 45 GQFDVSTQATTTATGQ--TAVATSGQPTGPTTTPAPASQ--VGQGTGAVD-- 99

Oy 73 RIEBKOLAATVGT--SVASISNNATRLPVV-----DQCVTLAG 114

Db 100 -----TTTATTTQATSAQAAGTQATPAVGCQAPATAPROGNKFTLSQ 151

Oy 115 FPSLSGNRIYCGGVSGQ-----VMDIVSPMGFVETSTHLSLI 158

Db 157 QSTGTGYN---FSLGTGSNSYVPGSYPMQPTAPSPFTSTSTOPTSDOSSY 208

Oy 159 SNPQWNASNNRCDCEFKRKRLDDQNNVRSNGSGSKYTIPIPMNGY---DOYLL 214

Db 209 SQONTGQPSST-----GQSSYQSSSTGQGPPTSPPTGSGYSAQPSQ 256

Oy 215 QSDHRRSQGFLDIHRAAASVASSSTINPYFNATNHTPME----- 259

Db 257 QSSSYQSSSYQFQIH-----PSSMVGYSQSGSGFSGPCNRSMSGPNRGRG 305

Oy 260 --RPGSYMEGNFPGSKYKFFPKYGER 289

Db 306 GFDGNGSGGGGGGCGM-----GSAGER 330

RESULT 2

US-09-214-564A-4

: Sequence 4, Application US/09214564A

: Patent No. 6150515

: APPLICANT: Sharp, Phillip A.

: APPLICANT: Zhou, Qiang

: TITLE OF INVENTION: Tat-SF: Cofactor For Stimulation Of Transcriptional

: REGULATION OF HIV-1 TAT

: FILING DATE: 1996-07-03

: CURRENT APPLICATION NUMBER: US/09/214,564A

: PRIOR FILING DATE: 1999-08-18

: PRIOR APPLICATION NUMBER: US 60/021,218

: PRIOR FILING DATE: 1996-12-13

: PRIOR APPLICATION NUMBER: US 60/033,152

: PRIOR FILING DATE: 1997-07-03

: SOFTWARE: FastSeq For Windows Version 3.0

: SEQ ID NO 4

: LENGTH: 656

: ORGANISM: Homo sapiens

US-09-214-564A-4

Query Match 5.78; Score 93; DB 4; Length 656;

Best Local Similarity 20.28; Pred. No. 0.5;

Matches 67; Conservative 34; Mismatches 127; Indels 104; Gaps 13;

Oy 16 GPNDLRLRVVWSSGEITFTTLKGRKP---GSKTQKQKKPPLRGKGVAKLEQ 72

Db 45 GQFDVSTQATTTATGQ--TAVATSGQPTGPTTTPAPASQ--VGQGTGAVD-- 99

Oy 73 RIEBKOLAATVGT--SVASISNNATRLPVV-----DQCVTLAG 114

Oy 115 FPSLSGNRIYCGGVSGQ-----VMDIVSPMGFVETSTHLSLI 158

Db 157 QSTGTGYN---FSLGTGSNSYVPGSYPMQPTAPSPFTSTSTOPTSDOSSY 208

Oy 159 SNPQWNASNNRCDCEFKRKRLDDQNNVRSNGSGSKYTIPIPMNGY---DOYLL 214

Db 209 SQONTGQPSST-----GQSSYQSSSTGQGPPTSPPTGSGYSAQPSQ 256

Oy 215 QSDHRRSQGFLDIHRAAASVASSSTINPYFNATNHTPME----- 259

Db 257 QSSSYQSSSYQFQIH-----PSSMVGYSQSGSGFSGPCNRSMSGPNRGRG 305

Oy 260 --RPGSYMEGNFPGSKYKFFPKYGER 289

Db 306 GFDGNGSGGGGGGCGM-----GSAGER 330

RESULT 3

US-08-891-640-2

: Sequence 2, Application US/08891640

: Patent No. 628173

: APPLICANT: Grunewald, Pierre

: APPLICANT: Gronemeyer, Hansrich

: APPLICANT: Voegel, Johannes

: TITLE OF INVENTION: Transcriptional Intermediary Factor-2

: NUMBER OF SEQUENCES: 14

: CORRESPONDENCE ADDRESS:

: ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.

: CITY: New York, New York Avenue, NW, Suite 600

: STATE: Washington

: COUNTRY: USA

: ZIP: 20005-3934

: COMPUTER: IBM PC compatible

: OPERATING SYSTEM: PC-DOS/MS-DOS

: RELEASE DATE: 1996-07-03

: CURRENT APPLICATION NUMBER: US/08/891,640

: FILING DATE: 1996-07-03

: CLASSIFICATION: 435

: APPLICATION NUMBER: US 60/021,247

: FILING DATE: 12-JUL-1996

: ATTORNEY/AGENT INFORMATION:

: NAME: Sterne, Eric K. 36 688

: REFERENCE/DOCKET NUMBER: 1393.0130001/EKS

: TELECOMMUNICATION INFORMATION:

: TELEPHONE: 202-371-2600

: INFORMATION FOR SEQ ID NO: 2:

: SEQUENCE CHARACTERISTICS:

: LENGTH: 1464 amino acids

: TYPE: amino acid

: MOLECULE TYPE: protein

US-08-891-640-2

Query Match 5.69; Score 92; DB 4; Length 1464;

Best Local Similarity 15.25; Pred. No. 2;

Matches 56; Conservative 45; Mismatches 107; Indels 104; Gaps 10;

Oy 13 NVENNDNLRL-----NTELVNWSGEITFTTLKGRPSKSTQCKKKPTLRG 63

Db 1211 SVNSVNLVRLRGVPTQATPAAGAAQKQRELLNGLHLEKQK---HQOQVQRTLRAG 1267

Oy 64 MGVALKRLQRIIEEKKLAATVGT--SVASISNNATRLPVV-----NATRLPVVQDVGVLGPF 117

118 SLGSRNRYGGVSGSQWIDYFISPKGFVETSTTHELSSINPQNASSNKCDCIFK 177
 1315 FTGAT-----TPGSPJMP-----: : : : :
 178 KKLGDGNNVNSGSGSKYTIMPPNKGQYDYLQSDHDSQGLYDRIARAASV 237
 1346 AYQASDINWAGNCGNENFSGSPHFG-----: : : : :
 238 SASSTIYNFKNATNFGMEFGSGYMEGNFNGSGVETFFPKYKGRVSVATRS 297
 1382 SKMSNNINNVSMATN-----: : : : :
 298 SLGSCSPWID 309
 1422 SLGSSNKPBOVN 1433

RESULT 4

US-09-185-160-13
 Sequence 13, Application US/09185160
 Patent No. 6252137
 GENERAL INFORMATION:
 APPLICANT: JOAN T.
 APPLICANT: HARDER, PATRICIA A.
 TITLE OF INVENTION: SOYBEAN HOMOLOG OF A SEED-SPECIFIC
 TITLE OF INVENTION: TRANSCRIPTION ACTIVATOR FROM PHASBOULS
 NUMBER OF INVENTIONS: 1
 NUMBER OF INVENTIONS: 1
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: E. I. DUPONT DE NEMOURS AND COMPANY
 STREET: 1007 MARKET STREET
 CITY: WILMINGTON
 STATE: DELAWARE
 COUNTRY: USA
 ZIP: 19898
 COMPUTER READABLE FORM:
 FILE NAME: 185-160-13 3.50 INCH
 COMPUTER: IBM PC COMPATIBLE
 OPERATING SYSTEM: MICROSOFT WINDOWS 95
 SOFTWARE: MICROSOFT WORD VERSION 7.0A
 CURRENT APPLICATION DATA:
 FILING DATE: 09/09/185,160
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 CHARGE NUMBER: 0/065 459
 FILING DATE: NOVEMBER 17, 1997
 ATTORNEY/AGENT INFORMATION:
 NAME: MANARIAN, WILLIAM R.
 REGISTRATION NUMBER: 41,173
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 302-992-4926
 TELEFAX: 302-773-0164
 INFORMATION FOR SEQ ID NO: 13:
 SOURCE: SOYBEAN
 LENGTH: 734 amino acids
 TYPE: amino acid
 STRANDNESS:
 TOPOLOGY: linear
 MOLECULE TYPE: peptide

US-09-185-160-13
 Query Match
 Best Local Similarity 21.8% Pred: No. 0.97
 Matches 70, Conservative 30, Mismatches 83, Indels 130, Gaps 16:

5 LEFSTQNSVGNPDLLRNLVHSSGEITRTLSKRGKPSQKQKPTLNG 63
 228 LEWLSKMSQV-SAND-LRAVL-----KATLISAAARJGGOREMKQKLLLEN 277
 64 MEVALKSPRIETEEKOLAAATVDTSTSVASISNNATVPVPPVDDGVVLQG-PRSELGSLN 122

278 VQTSHLNKKRKE-----NNGSSISS-----: : : : :
 123 RITCGVGSQWQWIDVISPMPGVETSTTHELSSINPQNASSNRCDCPKKKRLD 182
 312 -----NKTPLLSQTTAT-----DONLAWPQTPQWQYVDDYTPGASNNI 376
 183 GQNNVYNSGSGSKYTIMPPNKGQYDYLQSDHDSQGLYDRIARAASVSA 239
 330 --NKTPLLSQTTAT-----DONLAWPQTPQWQYVDDYTPGASNNI 376
 240 SST-----TINY-----FBNATNFGMEFGSGYMEGNFNGSGG 275
 277 SATNHNNSNPQTPQWQYDYLQSDHDSQGLYDRIARAASVSA 433
 376 -----VKTEFFPKYKGRV 290
 434 GSYGNHYTPFFHPG-GDKRL 453

RESULT 5

US-08-441-139-9
 Sequence 9, Application US/08441139
 Patent No. 6252137
 GENERAL INFORMATION:
 APPLICANT: MITTUP, Dr. Karl D.
 APPLICANT: ROBINSON, Anne S.
 TITLE OF INVENTION: METHODS FOR INCREASING SECRETION OF
 TITLE OF INVENTION: SECRETION OF CONTINUALLY EXPRESSED PROTEINS
 NUMBER OF INVENTIONS: 20
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: SCULLY, SCOTT, MORPHY & PRESSER
 STREET: 400 garden city plaza
 CITY: garden city
 STATE: NY
 COUNTRY: USA
 ZIP: 11530
 COMPUTER READABLE FORM:
 FILE NAME: 441-139-9 3.50 INCH
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 FILING DATE: 05/08/441,139
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 CHARGE NUMBER: US 08/089,997
 FILING DATE: 06-JUL-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: DIGILIO, Frank S.
 REGISTRATION NUMBER: 31,346
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 516-742-4343
 TELEFAX: 516-742-4366
 INFORMATION FOR SEQ ID NO: 9:
 SOURCE: HUMAN
 LENGTH: 1250 amino acids
 TYPE: amino acid
 STRANDNESS: linear
 MOLECULE TYPE: protein

US-08-441-139-9
 Query Match
 Best Local Similarity 15.9% Pred: No. 0.6
 Matches 81, Conservative 49, Mismatches 138, Indels 139, Gaps 17:

2 ATSLFPTMSTQNSVGNPDLLRNLVHSSGEITRTLSKRGKPSQKQKOKK 58
 453 ATSNWLTSPSSSSLSKDLVSRKASSTINDSLGSLSPKSNVRSLSKQPTQKK 512

QY 59 ---PTLRGNGVAKLERQIERKK-----OLAAATVG---DTSSVASIS 95
 Db 513 MDVPEVSGSLLVZEERIDMKYKIVAGHYAVLIDRIPOGLFOTGLLRPQSOANSON 572
 QY 97 WNKATPE---VVPDGVVGPSSLG-----SNELCYGGVSSQWIDPVI 140
 Db 573 NKPOESKIAWFKTDKVLIAITPELAPQDVANADKISLEKFLVASIKRW-----PIT 627
 QY 141 S---PHREFTVSTTHRESLSNPOPNASNNKCECFKKKLKEDDOONVNSGGGSEK 198
 Db 628 SLIMPGL---VSELAGDHPD-----DTEDISLRDN--NFLS 560
 QY 199 YTHMPPNNGVDOYLQSDHQRSGOFLVHRIARAAVSASASTINFPNATHFGM 258
 Db 661 -----NEYLDONKQEKSTOPLPLTAESLEYRNTDTNETNIFALSELOW 709
 QY 259 EEP-----QSYMEGRPNRSGCV----- 276
 Db 710 SEPALVRNNGTLELCHVVDVDSYHIERESSVBRKRESSAVMPQKVLNLPQSEN 769
 QY 277 KEYFFKPGTGVRSVAVT-----TSSLVGD--CSPNITDLSLK 313
 Db 770 DELSIWAGNSASTLSVYTLQSDSLRKSTWQSTWSTISPSNT-LSLE 815

RESULT 6

US-08-442-859-4

Sequence 4, Application US/0843104

Patent No. 5691162

GENERAL INFORMATION:

APPLICANT: Shuster, Jeffrey R.

APPLICANT: Madden, Mark

APPLICANT: Meyer, Donna L.

APPLICANT: Fuhsiang, Claus

APPLICANT: Branner, Sven

TITLE OF INVENTION: NOVEL METALLOPROTEASE HAVING INCREASED

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:

ADDRESSER: No. 5691162 No. 5691162disk of No. 5691162th America, Inc.

STREET: 405 Lexington Avenue, 64th Floor

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10174-6401

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.25

APPLICATION NUMBER: US/08/443,104

FILING DATE: 17-MAY-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/7398,489

FILING DATE: 03-MAR-1995

ATTORNEY/AGENT INFORMATION:

NAME: Agis Dr., Cheryl H.

REGISTRATION NUMBER: 34,086

REFERENCE/DOCKET NUMBER: 4180,010-US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-867-0123

TELEFAX: 212-878-9655

INVENTOR CHARACTERISTICS: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 632 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

Query Match 5.3%; Score 86; DB 1; Length 632;
 Best Local Similarity 22.94; Pred. No. 2.6;
 Matches 61; Conservative 39; Mismatches 104; Indels 62; Gaps 15;
 QY 44 GRKSGKGGQKQK-----KPLRGNGVAKLERQIERKKKOLAAATVGTDSVASIS 98
 Db 138 QKTPCLPLTKHSGQVDAKLDQVDVLSL-PVEADKAAKAKSKNHYT--TCTGTGVS 194
 QY 99 ATPLPVPDPVQVLQ--GFSSLSNRI--YGGVSGSV--NIDPV-----ISWPF 145
 Db 195 EAKLTLYVDNEMKELTKRWETDIDVNLITLYVNAKTDGVGVYVYATVYKVPW 254
 QY 146 VSTSTTHLSLSINPNFNAS-----SNKRCUTCFKKRLDQONVNSGGGFS 197
 Db 255 NDKSSGR--STVENPNKLASFTNLSGNSNYIT-----RSGNGIAQVPSGS 304
 QY 198 KYTIMPPMGVDOYLQSDHQRSGOFLVHRIARAAVSASASTINFPNETH 250
 Db 305 TV-----LNVYSP-----DSPSLAFETVISTITPTTTRASIAQLPYTAKT--HD 350
 QY 251 AYVMTGPMEEFGSYMEGRPNRSGCV 276
 Db 351 LYLQGLQTEQNGFQTN--NGQGV 374

RESULT 7

US-08-442-859-4

Sequence 4, Application US/0842859

Patent No. 5697729

GENERAL INFORMATION:

APPLICANT: Shuster, Jeffrey R.

APPLICANT: Madden, Mark

APPLICANT: Meyer, Donna L.

APPLICANT: Fuhsiang, Claus

APPLICANT: Branner, Sven

TITLE OF INVENTION: NOVEL METALLOPROTEASE HAVING INCREASED

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:

ADDRESSER: No. 5697729 No. 5697729disk of No. 5697729th America, Inc.

STREET: 405 Lexington Avenue, 64th Floor

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10174-6401

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.25

APPLICATION NUMBER: US/08/442,859

FILING DATE: 17-MAY-1995

CLASSIFICATION: 252

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/7398,489

FILING DATE: 03-MAR-1995

ATTORNEY/AGENT INFORMATION:

NAME: Agis Dr., Cheryl H.

REGISTRATION NUMBER: 34,086

REFERENCE/DOCKET NUMBER: 4180,010-US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-867-0123

TELEFAX: 212-878-9655

INVENTOR CHARACTERISTICS: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 632 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

Query Match 5.3%; Score 86; DB 1; Length 632;
 Best Local Similarity 22.9%; Pred. No. 2.6; Gaps 15;
 Matches 61; Conservative 39; Mismatches 104; Indels 62;

Y 44 GRPKSTGQOKOOR----KPTLRGKGVAKLERQRIEERKKOLAATVGTSSVASISNN 98
 DB 138 GRPOPIKDKKOPVDALQDYVLSL--PVEADKAKAKSKNHTP--TGTGTGSKP 194
 Y 99 ATFLRVDVDDVQLQ---GPPSLSGSNRI--YCGVGSGOV--MIDPV-----ISPMGF 145
 DB 195 EAKLYTVDENKEKLWTRVEDIDVDMALLTYVNAKTDVGVGVDTVNAATKYTPMGV 254
 Y 146 VESSSTHLSISNPFNFA--SNRKCDCFCFKKLDGDNVNRSGSGS 197
 DB 254 VESSSTHLSISNPFNFA--SNRKCDCFCFKKLDGDNVNRSGSGS 197
 Y 255 MDPSKGR--STVENPWLAASEFTWLSGDSNNYTT-----RGNKIAQVNPSSGS 304
 DB 198 KTYMTPPMPGVDYVLLQSHDHQSGQFLYDHRIA-----RAASVASASTTINPYNE 250
 Y 305 TY-----LNNRP-----DPSLKFEDYSTSTPTPTTFRDASIAQLPTANKY-HD 350
 DB 251 ATNHTGMEFSGSYMGKPNCSGV 276
 Y 251 ATNHTGMEFSGSYMGKPNCSGV 276
 DB 351 LLYLLGFTQAGNFTNN--NQSGV 374

ESUT 8 8-489-4
 Sequence 4, Application US/08198489

Patent No. 5843753
 GENERAL INFORMATION:
 APPLICANT: Shuster, Jeffrey R.
 INVENTOR: Shuster, Jeffrey R.
 APPLICANT: Meyer, Dennis L.
 APPLICANT: Fuglsang, Claus
 APPLICANT: Branner, Sven
 TITLE OF INVENTION: NOVEL METALLOPROTEASE HAVING INCREASED
 THERMAL STABILITY
 NUMBER OF SEQUENCES: 10
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: NO. 5843753rd of No. 5843753th America, Inc.
 ADDRESS: 405 Lexington Avenue, 64th Floor
 CITY: New York
 STATE: New York
 COUNTRY: USA
 ZIP: 10174-6401
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Tape
 OPERATING SYSTEM: IBM PC compatible
 SOFTWARE: PC-DOS/MS-DOS
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/398,489
 FILING DATE: 03-MAR-1993
 CLASSIFICATION: 435
 ANOTHER/AGENT INFORMATION:
 NAME: Aris dr., Cheryl H.
 REGISTRATION NUMBER: 34,086
 REFERENCE/DOCKET NUMBER: 4180,010-US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212-878-9653
 TELEFAX: 212-878-9653
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 632 amino acids
 TYPE: amino acids
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 PCT US95-05534-4

Query Match 5.3%; Score 86; DB 2; Length 632;
 Best Local Similarity 22.9%; Pred. No. 2.6;
 Matches 61; Conservative 39; Mismatches 104; Indels 62; Gaps 15;

Y 44 GRPKSTGQOKOOR----KPTLRGKGVAKLERQRIEERKKOLAATVGTSSVASISNN 98
 DB 138 GRPOPIKDKKOPVDALQDYVLSL--PVEADKAKAKSKNHTP--TGTGTGSKP 194
 Y 99 ATFLRVDVDDVQLQ---GPPSLSGSNRI--YCGVGSGOV--MIDPV-----ISPMGF 145
 DB 195 EAKLYTVDENKEKLWTRVEDIDVDMALLTYVNAKTDVGVGVDTVNAATKYTPMGV 254
 Y 146 VESSSTHLSISNPFNFA--SNRKCDCFCFKKLDGDNVNRSGSGS 197
 DB 254 VESSSTHLSISNPFNFA--SNRKCDCFCFKKLDGDNVNRSGSGS 197
 Y 255 MDPSKGR--STVENPWLAASEFTWLSGDSNNYTT-----RGNKIAQVNPSSGS 304
 DB 198 KTYMTPPMPGVDYVLLQSHDHQSGQFLYDHRIA-----RAASVASASTTINPYNE 250
 Y 305 TY-----LNNRP-----DPSLKFEDYSTSTPTPTTFRDASIAQLPTANKY-HD 350
 DB 251 ATNHTGMEFSGSYMGKPNCSGV 276
 Y 251 ATNHTGMEFSGSYMGKPNCSGV 276
 DB 351 LLYLLGFTQAGNFTNN--NQSGV 374

RESULT 9
 PCT US95-05534-4

Sequence 4, Application PC/TUS9505534
 GENERAL INFORMATION:
 APPLICANT:
 INVENTOR:
 TITLE OF INVENTION: NOVEL METALLOPROTEASE HAVING INCREASED
 THERMAL STABILITY
 NUMBER OF SEQUENCES: 10
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Novo Nordisk of North America, Inc.
 STREET: 405 Lexington Avenue, 64th Floor
 CITY: New York
 STATE: New York
 COUNTRY: USA
 ZIP: 10174-6401
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Tape
 OPERATING SYSTEM: IBM PC compatible
 SOFTWARE: PC-DOS/MS-DOS
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US95/05534
 FILING DATE: 03-MARCH-1995
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/238/108
 FILING DATE: 04-MAY-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/298,489
 FILING DATE: 03-MARCH-1995
 ANOTHER/AGENT INFORMATION:
 NAME: Aris dr., Cheryl H.
 REGISTRATION NUMBER: 34,086
 REFERENCE/DOCKET NUMBER: 4180,204-WO
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212-878-9653
 TELEFAX: 212-878-9655
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 632 amino acids
 TYPE: amino acids
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 PCT US95-05534-4

Query Match 5.3%; Score 86; DB 5; Length 632;
 Best Local Similarity 22.9%; Pred. No. 2.6;


```

EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,894
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/049,610
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,911
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,936
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,936
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,970
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,974
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,911
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,864
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,864
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,631
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,845
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,861
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,892
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/057,771
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/047,595
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/047,595
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/047,599
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/047,599
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/047,588
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/047,585
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/047,586
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/047,590
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/047,594
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/047,589
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/047,593
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/047,514
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/047,578
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,576
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/047,501
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/043,570
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/056,632
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,664
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,976
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,881
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,909
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,975
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,862
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,887
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,908
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/048,564
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/057,650
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/056,884
EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/057,669
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/049,610
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/061,060
; EARLIER FILING DATE: 1997-10-02

Query Match
Score 85.5; DB 4; Length 349;
Predicted Similarity 22.5%; Predicted No. 1.6;
Matches 30; Conservative 30; Mismatches 2; 66; Indels 25; Gaps 3;

OY 53 QQRKQPLKMGVAKLEQRLEEKQLAAATVGDTS-----SVASISNNATRLPVP 105
DB 138 QEMENKTYKDDGATPIKRNVSDEHYVSCISDKWTETVLPFTSDNGIROSSI 197
OY 106 VDPGVGLQFPSSISGNNYICGGVSGQWMDIPVSPGVFTSTHLESSISNPOFN 165
DB 198 IURPTE-QULPSPNSIS-----VETRMVYPSFSEDMNSIRSHAAE 239
OY 166 ASSNNKUTCTFKKRLGQDNQNVRSNGGF 196
DB 240 CSNNGRYDCKEFPNOLKCSKOSTLAEESZF 270

RESULT 11
US-09-071-035-484
; Sequence 484; Application US/09071035
; Title of Invention: Enterococcus faecalis polynucleotides and polypeptides
; Applicant: Gil H. Choi
; Number of Sequences: 496
; Number of Inventions: 1
; Address: Human Genome Sciences, Inc.
; Street: 9410 Key West Avenue
; City: Rockville
; State: Maryland
; Country: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; OPERATING SYSTEM: MS-DOS Version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/071,035
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY AGENT INFORMATION:
; NAME: A. Anders Brooks
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB369P2
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 484:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 349
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-071-035-484
Score 85.5; DB 4; Length 422;
Best Local Similarity 22.5%; Pred. No. 1.6;
Matches 62; Conservative 31; Mismatches 119; Indels 63; Gaps 9;

OY 49 SKTGQKQKPLKMGVAKLEQRLEEKQLAA-----ATVGVSSVASISNNATRLPVP 104
DB 138 QEMENKTYKDDGATPIKRNVSDEHYVSCISDKWTETVLPFTSDNGIROSSI 197

```


y 184 DQNNVRSNG-----GGFSKYTMIPPPNPKYDVLQLSDIHQ 220
 |||
 b 460 INHIGQGNQAQALUOGLMTSLINGLPAITLSOMMYQAMPPLANQPIIMQAQMGO 519
 |||
 y 221 RSQGF-LDYCHRAAAKSVASSTINYPE 248
 |||
 b 520 QQNQLQHQQMOOHNSSTTSHINSFP 548
 |||
 RESULT 14
 S-08-346-128-35
 Sequence 35, Application US/08346128
 GENERAL INFORMATION
 Patent No. 5856441
 Invention Title: Tackoag, Soyidin et al.
 TITLE OF INVENTION: Human No. 5856441ch And Delta, Binding Domains
 IN TOPOTHMIC PROTEINS, AND METHODS BASED THEREON
 NUMBER OF SEQUENCES: 37
 CORRESPONDENCE ADDRESS: 37
 STREET: 1155 Avenue of the Americas
 CITY: New York
 STATE: New York
 COUNTRY: U.S.A.
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 CURRENT APPLICATION DATA: Release #1.0, Version #1.25
 CURRENT APPLICATION NUMBER: US/08/346,128
 FILING DATE:
 CLASSIFICATION: 530
 PRIORITY APPLICATION DATA:
 PRIORITY APPLICATION NUMBER:
 FILING DATE: 30-APR-1992
 ATTORNEY/AGENT INFORMATION:
 NAME: MISTOCK, S. Leslie
 REGISTRATION NUMBER: 8,979
 RESIDENCE NUMBER: 8,979
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212 790-9090
 TELEFAX: 212 869864/9741
 TELEX: 661413 PENNE
 MAILING ADDRESS:
 INVOICE CHARACTERISTICS: 35;
 LENGTH: 654 amino acids
 TYPE: amino acid
 STRANDNESS: single
 MOLECULAR TYPE: protein

[illegible]

Query Match	5.24;	Scop: 85;	DB 3;	Length 454;
Best Local Similarity	18.64;	Prod No. 3.5;	35;	
Matches	50;	Conservative 40;	Mismatches 103;	Indels 76; Gaps 9;
Qy	46	KPGKGTGQOKRPTLGMVQAKLEQRTEERKQALAAATVCTSTSVASIS	-----	96
Dy	290	KISVGS-----KAAKPSGKQKELKARAKKQKDKTITLGGSSGVSPVSGSLSTH	347	
Qy	97	-----NNATLEPVDEVTQLQCFP-----SLGSGVATIDVP	139	
Dy	348	GILSDVSSPPLTSTPFQSPSPMLKHLSTPESQGLNKHINATIKQMAAGSNMADFAM	407	
Qy	140	ISMGVCTSTSTHELSSTNQMFAS-----SNGKQCFPK-----	183	
Dy	408	V-----PRULHNASSPTNINSNMHPTVGGAPTNISCCMLKRLRGVQVQYDP	459	
Qy	184	DQNNVRSNG-----GQFKTFTVTPMKGVYQDTLQSDHQ	220	
Dy	221	ESQGLQGNQMOQAAQNGHSTSLNGLPATLSNMTQMPNTFLANPFLKMQQMG	519	
Qy	221	ESQGLQGLQGNQMOQAAQNGHSTSLNGLPATLSNMTQMPNTFLANPFLKMQQMG	519	

Db 520 000N201105000HINSSTTTSTRINSPT 548

Search completed: January 26, 2003, 00:07:17
Job time : 28 secs

1	99.5	6.1	676	10	US-09-801-368-302	Sequence 302, App
2	95.5	5.8	625	10	US-09-801-368-294	Sequence 294, App
3	92.5	5.7	1239	12	US-09-801-368-577	Sequence 577, App
4	92.5	5.6	1484	12	US-09-801-368-286	Sequence 286, App
5	92.5	5.6	1484	12	US-09-801-368-286	Sequence 286, App
6	89.5	5.4	577	10	US-09-801-368-109	Sequence 109, App
7	89.5	5.4	577	10	US-09-801-368-109	Sequence 109, App
8	86.5	5.3	660	12	US-09-801-368-139	Sequence 139, App
9	86.5	5.3	1250	10	US-09-801-368-364	Sequence 364, App
10	84.5	5.2	3335	10	US-09-801-368-615-64	Sequence 64, App
11	84.5	5.2	3335	10	US-09-801-368-615-64	Sequence 64, App
12	84.5	5.2	3335	10	US-09-801-368-615-64	Sequence 64, App
13	83.5	5.1	1016	10	US-09-815-402-304	Sequence 304, App
14	83.5	5.1	1016	10	US-09-815-402-304	Sequence 304, App
15	82.5	5.0	382	10	US-09-825-301-449	Sequence 449, App
16	82.5	5.0	359	10	US-09-820-721A-1	Sequence 1, App
17	82.5	5.0	359	10	US-09-820-721A-1	Sequence 1, App
18	82.5	5.0	1518	10	US-09-801-368-174	Sequence 174, App
19	82.5	5.0	1703	10	US-09-801-368-340	Sequence 340, App
20	82.5	5.0	1703	10	US-09-801-368-340	Sequence 340, App

APPLICANT: PLOMAN, GREGORY
 APPLICANT: WYTE, DAVID
 APPLICANT: CAZNEPEL, SEAN
 APPLICANT: CHARYCZAK, GLEN
 APPLICANT: MANNING, GERRARD
 APPLICANT: PLOMAN, GREGORY
 TITLE OF INVENTION: NOVEL PROTEASES
 FILE REFERENCE: 038602/1214
 CURRENT APPLICATION NUMBER: US/09/888,615
 PRIOR APPLICATION NUMBER: 03/01,047
 PRIOR FILING DATE: 2000-06-26
 PRIOR FILING DATE: 2000-06-26
 NUMBER OF SEQ ID NOS: 150
 SOFTWARE: Patent In Ver. 2.1
 SEQ ID NO 1
 LENGTH: 3343
 TYPE: PRT
 ORGANISM: Homo sapiens
 EXTRACT: W04.BES
 LOCATION: (1991)
 OTHER INFORMATION: Any amino acid
 J8-09-888-615-64

Query Match 5.24; Score 85.5; DB 10; Length 3353;
 Best Local Similarity 15.94; Pred. No. 50;
 Matches 30; Conservative 66; Indels 25; Gaps 3;

Y 53 QVKKQKPTLRKGVKALRQVTEFEKLAANVGVTS
 Y 3143 QKAKERTKDDCATPKRRVSDREHTEVTCISDMKTEFEVLPTSTSONETIHDSI: 3202
 Y 106 VQGVVLQFGSGSNRIYCGGVSGCWIDPVIWPFVETSTHLSISNQPNRY 165
 Y 3203 IDGTE-QDLSPENS-----VKREYMEVSSPEMSNITSCHASE 3244
 Y 165 ASSNNRCDTPFKRRKLDQNNVYNSGGF 196
 Y 3245 QSNKGVYDCKEFLKSGKSTLAKKESEP 3275

RESULT 10
 US-09-754-4
 Sequence 4, Application US/09855754
 Publication No. US20020192237A1
 GENERAL INFORMATION:
 APPLICANT: BOURSAUX, EDDIE
 APPLICANT: BOURSAUX, EDDIE
 TITLE OF INVENTION: POLYPEPTIDES CONTAINING POLYMORPHISMS OF THE REPEATED
 TITLE OF INVENTION: REGIONS OF PECTACTIN IN BORDETELLA PERTUSSIS,
 TITLE OF INVENTION: BORDETELLA PARAPERTUSSIS, AND BORDETELLA
 TITLE OF INVENTION: BRONCHISEPTICA, THEIR USE IN DIAGNOSTICS, AND IN
 COMPOSITIONS
 FILE REFERENCE: 03495-0206-00000
 CURRENT APPLICATION NUMBER: US/09/855,754
 CURRENT FILING DATE: 2001-09-10
 PRIOR APPLICATION NUMBER: 03/005,969
 NUMBER OF SEQ ID NOS: 24
 SOFTWARE: Patent In Ver. 2.1
 SEQ ID NO 4
 TYPE: PRT
 ORGANISM: Bordetella bronchiseptica
 J8-09-855-754-4

Query Match 5.24; Score 85; DB 9; Length 911;
 Best Local Similarity 19.94; Pred. No. 8.3;
 Matches 61; Conservative 31; Mismatch 98; Indels 116; Gaps 12;

Y 17 KPNLDLNLRYVNSGEIRITFKLRGFGSKT-----51
 Y 83 NPAEALFQNGSVYSQGLDFGTVR---RLFGTVTVKAKLVADIATLANVSDTRDDGT 139

QY 52 -----GQKKK-----PTLRGVKALKER-----QTEFEKALQIA 82
 DB 140 ALTYAGQNAQSTADTQAGAGVYRVRGANTVORSTVVGGLGTEIGTLPQEDLPSP 199
 QY 83 AAHVQGVSYVASISNNAETVPYVQVVLQGFPSLGSNRY-----YGVGVSQVMI 136
 DB 200 KVLADTIS-----VTAVPASCAPAAV-----SVFGANELYDGGHITGGRAVAM 246
 QY 137 D-----VYSPIMGVVFSTFHIEL-----SSISNQ 162
 DB 247 DQAIVLHORATIRGADPAGGAVPGAGVGGPGLDGGVYDVSUSTVAQSLVYEPQ 306
 QY 163 MFNASSNNP-CQTFKKRRKLDQNNVYNSGGFSKYTIMPPMPCNTDQYLASDHOR 221
 DB 307 LCAATPAGGAVRVYSSGSLSNPHQVETEGCA-----HREPPASPL-SITLACAAQ 361
 QY 222 SQGFY 227
 DB 362 GRALLT 367
 RESULT 11
 US-09-741-669-304
 Patent No. US20030022718A1
 GENERAL INFORMATION:
 APPLICANT: Forsyth, R. Allyp
 APPLICANT: Olesen, Kristian
 APPLICANT: Olesen, Kristian
 TITLE OF INVENTION: Genes identified as required for
 FILE REFERENCE: ELITRA.009A.US/09/741,669
 CURRENT FILING DATE: 2000-12-19
 PRIOR APPLICATION NUMBER: US 60/173005
 PRIOR FILING DATE: 1998-12-23
 NUMBER OF SEQ ID NOS: 481
 SOFTWARE: Patent In Ver. 2.1
 SEQ ID NO 304
 LENGTH: 1325
 TYPE: PRT
 ORGANISM: Escherichia coli
 US-09-741-669-304

Query Match 5.24; Score 84.5; DB 10; Length 1325;
 Best Local Similarity 22.84; Pred. No. 16;
 Matches 84; Conservative 41; Mismatch 145; Indels 99; Gaps 21;

QY 8 NSTQDSNVNPN-DLLNRTLVNSGEIR-ETLKSAG-----RKPG 48
 DB 739 MGSSEVTINPHDILAST-----NSAGDTLTNALKGGVLMYVQLSSSKRMFGTHYG 794
 QY 49 SK-----TQCKKXKPTLRKGVKALRQVTEFEKLAANVGVTSVASISNA-----99
 DB 795 TEFAVQAQKDSPTFLERONTAALTANLQSDSENTSYKVVG-OSIGCLAHNGKTIIFD 853
 QY 100 TELP-----VYDQGVVLQFGPSLSGNRIYCGGVSGCWIDPVIWPFVETSTHLS 149
 DB 854 TDTPAATLAEGYISVDTLVVGAGDYTKGRNVQV---NGTGQVLID-VPEKNDPMANNP 909
 QY 150 STT-----HEUSISINPQNFVASSNNRCDTPFKRRKLDQ-----QNNVYNSG 193
 DB 910 LITLKLLEHSDNP-GVQLVKAQTVYSSGSLSTLRQLQDQVDAKTLHINQNTVADG 968
 QY 194 G-GSKYTIWPPNNG-YDQYTLQAS-DHIOHSQGFVYDIRIARAASVASS-----TT 243
 DB 969 DYQRELT-----APNGVLVYVNCJLAKNTUGCKLTIARIGGAGATKNAUGGDLA 1025
 QY 244 TN-----PYFENATINCPMEERFGSYMEINPNG-----SGGVKEVFP 281
 DB 1026 INTVQVSLNSWQNDYGGATVYQMGTEFLTDADGALONTRELATSNMAIVLQNGSTOYTV 1085

CURRENT FILING DATE: 2001-08-10
 PRIOR APPLICATION NUMBER: PCT/US90/05882
 PRIOR FILING DATE: 2000-03-08 124, 270
 PRIOR PUBLICATION NO.: 5,812,812
 PRIOR FILING DATE: 1999-03-12
 NUMBER OF SEQ ID NOS: 1634
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 1409
 TYPE: PRF
 ORGANISM: Homo sapiens
 S-09-925-301-1409

Query Match 5.14; Score 83; DB 10; Length 488;
 Best Local Similarity 26.98; Pred. No. 5.2;
 Matches 42; Conservative 13; Mismatches 53; Indels 48; Gaps 8;
 y 9 STONKSGNDLLAN-----TRLVNNSGQIFRTELKSGRKPFSGTCQKQKKEPLR 62
 |||||
 b 338 STONMKGKATPSANOKPFGVLGLATLSKGGI-----GSKIGNSNTTPTV- 384
 y 63 GHWAKLERQREERKGLAAVTVGUTSVASIS--NNATRLPVVPDGVVLQGFPSGLS 121
 |||||
 b 385 -----PLKPPPLLTGKTLGSLNSVDNKKVGLP-SFSSLVGSSQLSG 429
 y 122 NRIYGGVGGQVMDPVISMPGFVETSTHIELSS 137
 |||||
 b 430 N-----GNGGTS-----GPGSAST-INTTSSSS 452

RESULT 35
 S-09-925-301-368-210
 Sequence 210: Application US/09801368
 Patent No. US20020128250A1
 GENERAL INFORMATION:
 APPLICANT: Bushy, Robert
 APPLICANT: Bushy, Robert
 APPLICANT: Hecht, Peter
 APPLICANT: Holtzman, Doug
 APPLICANT: Madden, Kevin
 APPLICANT: Maxon, Mary
 APPLICANT: Nishida, Kenji
 APPLICANT: No. US20020128250A1man, Thea
 APPLICANT: Royer, John
 APPLICANT: Selama, Sofie
 APPLICANT: Sherman, Amir
 APPLICANT: Summers, Eric
 TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
 FILE REFERENCE: 109272.147 US 09/801.368
 CURRENT FILING DATE: 2001-03-07
 PRIOR APPLICATION NUMBER: US 09/487.558
 PRIOR FILING DATE: 2000-01-19
 PRIOR APPLICATION NUMBER: US 60/160.587
 PRIOR APPLICATION NUMBER: US 60/160.587
 NUMBER OF SEQ ID NOS: 1634
 SOFTWARE: PatentIn version 3.0
 SEQ ID NO 210
 LENGTH: 382
 TYPE: PRF
 ORGANISM: Saccharomyces cerevisiae

S-09-801-368-210
 Query Match 5.04; Score 82.5; DB 10; Length 382;
 Best Local Similarity 22.34; Pred. No. 4.4;
 Matches 69; Conservative 39; Mismatches 98; Indels 101; Gaps 16;
 y 35 IRETELKSGRKP-----SKTSQKQKPPPLGNGVAKLERQREERKGLAA 83
 |||||
 b 36 LNNIT-----GEPHICAFPCCKSFSSSDELKNNHRTTGTGSGSRUKANSGVGE---f 87
 y 84 ATVGDTSSVASISNNATRLPVVPDGVVL-QGFPSSLSGN-RIYCGVGSGCV----- 134

Db 88 UTYSCTPIITAS-----GVMHQPIPOVLPAANALVQVAMNGNIIHAPNAV 113
 ||:::|
 Qy 135 --WIDPV-----ISMPGEVETISSTHELSSINQPMENASSNN 170
 ||:
 Db 134 HPWVPIPAQAPAPHASASPOATSPMP1STVTPVPSQFTSFQSSGSGTO-----SNS 188
 ||:::|
 Qy 171 KQTCFCKKRLGQDNVNSKGGGFSKYTMID-PMNCKY--DQYLQSDHQRSS----- 222
 ||:
 Db 189 DVSSIFSNKV-----RVNTPRS-----VNSPNDCLHQQRPOQVQIQTASPSVA 235
 ||:
 Qy 223 QGFIYDHRITARRAS-----VSASSTTI-NFYFNATNHTGPMERGSWEGNPNRNS 273
 ||:
 Db 236 KQKMTFANSLASALSTLQKTPNAPSTTIESPSSDSNHSASSAHLSPFSNAPSOL 295
 |||||
 Qy 274 GQKYE 280
 |||||
 Db 296 AVARELE 302

Search completed: January 26, 2003, 00:07:52
 Job time : 23 secs

This Page Blank (uspio)

A; Introns: 10/2; 37/2; 171/3; 206/1; 285/3; 662/2; 738/3
A; Note: F22K18.120

Query Match	6.0%	Score 98	D8 2	Length 1480
Best Local Similarity	21.9%	Pred. No. 15		
Matches	80	Conservative	37	Mismatches 106; Indels 142; Gaps 19
Q#	3	TSLEFNSDQ	-----NSGNP-----	WELANLVY-VNKGSE 34
Q#	658	TSIFRPETPQ	QGRADPOTKRIWSEGDAAWKITVWSSGSSHTLIATITSEFRVWVDUS 717	
Q#	35	ITPLTKSKRPIS	-----KTKQKQKPK-----	TLRGMGVKYLE 70
Db	718	LOTESI	-----RPGSISADPKDMSWELARQDQKQFEERADQRAALAKLE 773	
Q#	71	---RRTIEE---	KQLGAAATVGDTSVASISNNATRLPDPVPOVGLQFFSFLGSNRI 124	
Q#	774	ELANRSDITEDSVSNKDA	-----SNSSPADPDP-----	GSUSNATNSV 816
Q#	125	YCGGSGSGVMDIPVISEPQVETISVTHLEISISPNFWMASNRQDTCFKKRLDQ 184		
Db	817	EPYG-GSKK	-----MTQRTSTETANNVGFSD 844	
Q#	185	NNNVYSSNGGFSKTVIPPNKGYDYLQSDHQSQDGLYDHRILARASVASTTI 244		
Q#	845	QNDLPED--GGASQKRL	-----GYKQ-----	KONTIEPKPKGSGVAVAEVD 867
Db	245	NP---VFYNE---	ATNTHGPRGZSGHGPQ--RNGSGVYVEFPFKKRGVRSVAVT 296	
Q#	888	VSPFVPEWVSSNSNSONDEATISVSAESTFKPKKRNKKKKK	-----AETATMTT 941	
Q#	297	SSZVG 301		
Db	942	RVANG 946		
<p> REQUEST 9 DESCRIPTION hypothetical protein T19D12.2 - <i>Caenorhabditis elegans</i> CSpecies: <i>Caenorhabditis elegans</i> CDate: 99-Oct-1999 sequence_revision 29-Oct-1999 <i>text_change</i> 29-Oct-1999 CAccession: T19D12.2 CAccession: T19D12.2 submitted to the EMBL Data Library, November 1995 A:Description: The sequence of <i>C. elegans</i> cosmid T19D12. A:Reference number: Z21513 A:Accession: T194368 A:Accession: T194368 A:Molecule type: DNA A:Residues: 1-561 <FAP> A:Cross-references: EMBL:U01263; FIDN:AC24423.1; GSPDB:GN00020; CESP:T19D12.2 A:Experimental source: strain Bristol N2; clone T19D12 A:Gene: CESP-T19D12.2 A:Map position: 2 A:Annotations: 251/1; 151/2; 214/1; 282/3; 333/2; 443/2; 469/2; 492/2 </p>				
Query Match	6.0%	Score 97.5	DB 2	Length 561
Best Local Similarity	21.8%	Pred. No. 4		
Matches	70	Conservative	42	Mismatches 108; Indels 101; Gaps 19

[illegible]

R; Nelson, J.; White, S.; Hawkins, J.; Wohldmann, P.
submitted to the EMBL Data Library, December 1997
A; Description: The sequence of *C. elegans* cosmid C02B10

A:Accession: J32554
A:Status: preliminary
A:Source: translated from GB/EMBL/DBJ
A:Size: 1168
A:Residues: 168
A:Cross-references: EMBL:AF03605; PIR:IA99200.1; GSPDB:GN00022; CESP:C02B10.5
A:Experimental source: strain Bristol N2; clone C02B10
A:Genetics:
A:Mutations: CESPDB: C02B10.5
A:Map position: 4
A:Introns: 61/3; 102/2; 188/3; 349/2; 641/1

RESULT 11
A56783
unknown protein F2H5.2 [unported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 sequence_revision 02-Mar-2001 #ext_change 31-Mar-2001
C:Accession: A56783
C:Author: Scher, J.B., Palm, C.J., Federspiel, N.A., Kaul, S., White, O., Altshuler, C.M., Chung, W.K., Cona, T., Conway, A.S., Conway, K.R., Creevy, T.H., Dewar, R.N., Hughes, B., Huizart, L.,

[illegible]

[illegible]

FT	DOMAIN	256	285	RNA-BINDING (RGM).
FT	DOMAIN	360	446	RANBP2-TYPE.
FT	2A1_FING	517	548	31 X APPROXIMATE TANDEM REPEATS.
FT	DOMAIN	8	285	
FT	REPEAT	17	27	1.
FT	REPEAT	28	34	2.
FT	REPEAT	35	42	3.
FT	REPEAT	43	50	4.
FT	REPEAT	51	59	5.
FT	REPEAT	60	68	6.
FT	REPEAT	69	75	7.
FT	REPEAT	76	84	8.
FT	REPEAT	84	90	9.
FT	REPEAT	92	110	10.
FT	REPEAT	111	116	11.
FT	REPEAT	117	125	12.
FT	REPEAT	126	136	13.
FT	REPEAT	135	146	14.
FT	REPEAT	144	156	15.
FT	REPEAT	164	170	16.
FT	REPEAT	171	177	17.
FT	REPEAT	178	188	18.
FT	REPEAT	190	196	19.
FT	REPEAT	199	201	20.
FT	REPEAT	202	206	21.
FT	REPEAT	207	212	22.
FT	REPEAT	213	218	23.
FT	REPEAT	219	224	24.
FT	REPEAT	225	230	25.
FT	REPEAT	231	238	26.
FT	REPEAT	239	245	27.
FT	REPEAT	246	252	28.
FT	REPEAT	253	260	29.
FT	REPEAT	260	276	30.
FT	REPEAT	277	285	31.
FT	DOMAIN	300	339	ANG/ELV/PRO-RICH.
FT	DOMAIN	453	532	ANG/ELV/PRO-RICH.
FT	MOD_RES	300	300	METHYLATION (D-).
FT	MOD_RES	302	302	METHYLATION (D-.) (BY SIMILARITY).
FT	MOD_RES	304	304	METHYLATION (D-.) (BY SIMILARITY).
FT	MOD_RES	319	319	METHYLATION (D-.) (BY SIMILARITY).
FT	MOD_RES	317	317	METHYLATION (D-.) (BY SIMILARITY).
FT	MOD_RES	321	321	METHYLATION (D-.) (BY SIMILARITY).
FT	MOD_RES	434	434	METHYLATION (D-.) (BY SIMILARITY).
FT	MOD_RES	454	454	METHYLATION (MONO- OR D-.) (BY SIMILARITY).
FT	MOD_RES	470	470	SIMILARITY.
FT	MOD_RES	489	489	METHYLATION (D-.) (BY SIMILARITY).
FT	MOD_RES	493	493	METHYLATION (D-.) (BY SIMILARITY).
FT	MOD_RES	502	502	METHYLATION (D-.) (BY SIMILARITY).
FT	MOD_RES	505	505	METHYLATION (D-.) (BY SIMILARITY).
FT	MOD_RES	562	562	METHYLATION (D-.) (BY SIMILARITY).
FT	MOD_RES	574	574	METHYLATION (D-.) (BY SIMILARITY).
FT	MOD_RES	574	574	METHYLATION (D-.) (BY SIMILARITY).
FT	MOD_RES	580	580	METHYLATION (D-.) (BY SIMILARITY).
FT	MOD_RES	588	588	METHYLATION (D-.) (BY SIMILARITY).
FT	MOD_RES	595	595	METHYLATION (D-.) (BY SIMILARITY).
FT	MOD_RES	599	599	METHYLATION (D-.) (BY SIMILARITY).
FT	MOD_RES	602	602	METHYLATION (D-.) (BY SIMILARITY).
FT	MOD_RES	606	606	METHYLATION (D-.) (BY SIMILARITY).
FT	MOD_RES	632	632	METHYLATION (D-.) (BY SIMILARITY).
FT	MOD_RES	635	635	METHYLATION (D-.) (BY SIMILARITY).
FT	MOD_RES	266	266	PHOSPHORYLATION (BY PKC) (BY SIMILARITY).
FT	SEQUENCE	655	AA	68418 MW: 50735ED5427049 CRC64:

Query watch

5.7a: Score 93: DB 1: Length 655:

Best Local Similarity

21.8a:

Pred. watch

Pred. No. 5.8:

Query Match	5.7%	Score 93;	DB 1;	Length 655;
Best Local Similarity	21.8%	Pred. No. 5.8;		

662 STTSLSLNGPKSLSS-----NGVSTVKPPT-KLSGTSTAPATVKVGLGKAT-----ITNK 911
 59 PTLGKGVAKLQRIEKEKULAAATGVSSVSSINNAATLEVPDPVGVVLQGPSS 118
 913 PTATSGASTDQVTRTLRLPLVSYNRRPA--TSGTGSVASSTARRQVFNKAG---SAPGS 966
 118 LSGNRITCGGVSGGVQDVIVQVSNQVYVTSSTHRLSSSLG-----OMFNASN 169
 967 AASTKVPAMTAPVK-PKVLSPSTISSTITVKKVPSSTSTSTSTSPKQKSNCLGK 1025
 170 NCUTCFKKRUDGQNNVRS-----NGGFSKTYMPPNPNQVQVLQSD 217
 1026 WSSSTTSATATITSTFARSAPKTHASLTAVNGKSSRLVLP-----GSSSTTTTSS 1081
 218 HHQSGSGELVRIIARAASVSSASTINPTFNPAINTGP-MEEFGSTMEGPNKNSGG 275
 1082 LKSSP-----LRAKCKKQKVLPOSKQATSSPVALAHSITLNG---GSDV 1129
 276 VKVEYFQPKYGE 288
 1130 APSNCEVPTINGQ 1142

RESULT 15
 NC02_HUMAN STANDARD: PRF: 1464 MA.
 NC02_HUMAN
 15599
 16-OCT-2001 (Rel. 40, Created)
 16-OCT-2001 (Rel. 40, Last sequence update)
 16-OCT-2001 (Rel. 40, Last annotation update)
 2182 TRANSFERRIN RECEPTOR COGNATE 2 (Transferrin receptor cognate factor 2)
 NC02 OR TF2.
 Homo sapiens (Human).
 Eukaryota, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Primates; Catarrhini; Hominoidea; Homo.
 NCBI_TaxId:9606;
 [1]
 SEQUENCE FROM N.A.
 TISSUE=Placenta;
 TISSUE_PUBMED=16470870;
 TISSUE_AUTHOR=Heine M.J.S., Zechel C., Chambon P., Gronemeyer H.;
 "TF2, a 160 kDa transcriptional mediator for the ligand-dependent
 activation function AF-2 of nuclear receptors.";
 EMBO J. 15:3667-3675(1996).
 CHARACTERIZATION.
 MEDLINE=98043472; PUBMED=9430642;
 Voegele J. J., Heine M.J.S., Tini M., Vivet V., Chambon P.,
 Gronemeyer H.; TF2 contains three nuclear receptor-binding motifs
 and mediates transactivation through CBP binding-dependent and
 -independent pathways.;
 EMBO J. 17:507-519(1998);
 FUNCTION=TRANSCRIPTIONAL COACTIVATOR FOR SPERMIO, RECEPTORS AND
 -1- FUNCTION=TRANSFORMATION OF THE SPERMIO BINDING DOMAIN
 (AF-2) BUT NOT OF THE MODULATING N-TERMINAL DOMAIN (AF-1).
 -1- SUBCELLULAR LOCATION: Nuclear.
 -1- SIMILILARITY: CONTAINS 1 PAS (PER-ARNT-SIM) DIMERIZATION DOMAIN.
 THIS SWISS-PROT entry is copyright. It is produced through a collabora-
 tion between the Swiss Institute of Bioinformatics and the EMBL outstation
 at the European Bioinformatics Institute. There are no restrictions on
 use by non-profit institutions as long as its content is in no
 way modified and this statement is included in the header of any
 redistributed file. For more information on the EMBL database policy
 or send an email to license@ebi.slb.ch.

EMBL: X27574; CAA62453.1;
 NC02_HUMAN
 GENBANK: F7169; NC02.
 MIM: 601993;

```

DR InterPro: IP001092; HUB basic.
DR InterPro: IP001610; PAC
DR InterPro: IP000014; PAS_domain.
DR Pfam: PF00989; PAS; 1.
DR SMART: SM00351; HUB; 1.
DR SMART: SM00352; PAC; 1.
DR SMART: SM00353; PAS; 1.
DR SMART: SM00354; PAS; 1.
DR PROSITE: PS50888; HUB2; 1.
DR PROSITE: PS50112; PAS; 1.
KW Transcription regulation; Activator; Nuclear protein.
FT DOMAIN 1254 1260 POLY-GLN
FT SEQUENCE 1464 AA; 159156 MW; 0AGIAA5D187804B CRC64;

Query Match
BLAST Local Similarity 19.5%
Matches 52; Conservative 36; Mismatches 88; Indels 85; Gaps 9;

QY 13 NSVGNNDLLR-----NTELVNSSGEIRTELKSGRKGSGKTKGQKKKPLRG 63
DB 1211 NSVSNVNTLRGVPTQPIAAGK-AQORELLNGLAQGM--HCOQVQGYTLAKRG 1267
QY 64 MGVALEHRIEEKKLAATVGDIVSSVASISN-----NATRLPVVDGVVLQGFPS 117
DB 1268 QGL-----NTPSWAPSGAPATSNRPIDPANAQQPPFPTGTSQQDFG 1314
QY 118 SUGSNRYTCGVGSGQWIDVISPMPGVETSTFHELSISNPQMFNASSNNRCDCEK 177
DB 1315 FTGAT-----TFUSTLGGP-----PMATQSPMHOQGAN-----P 1345
QY 178 KRLGDONNVNENSGESKTYTLPMPANETDYQLQSDHQRSGGFLYHRLAASV 237
DB 1346 AYOAPSDINGMAQGMGNSNFSQSQPPHG-----QQANT 1381
QY 238 SASSTTHYNENATNTGDM 258
DB 1382 SMYSNNAINVSMATN-TGGM 1401

```

Search completed: January 26, 2003, 00:04:15
 Job time : 34 secs